

Review

Trends in Using IoT with Machine Learning in Health Prediction System

Amani Aldahiri ^{1,*}, Bashair Alrashed ¹  and Walayat Hussain ² 

¹ Department of Cybersecurity, College of Computing Science and Engineering, University of Jeddah, Jeddah 21589, Saudi Arabia; baalrashed@uj.edu.sa

² School of Information, Systems and Modelling, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW 2007, Australia; Walayat.hussain@uts.edu.au

* Correspondence: ahal-dhahery@uj.edu.sa; Tel.: +966-556-331-212

Abstract: Machine learning (ML) is a powerful tool that delivers insights hidden in Internet of Things (IoT) data. These hybrid technologies work smartly to improve the decision-making process in different areas such as education, security, business, and the healthcare industry. ML empowers the IoT to demystify hidden patterns in bulk data for optimal prediction and recommendation systems. Healthcare has embraced IoT and ML so that automated machines make medical records, predict disease diagnoses, and, most importantly, conduct real-time monitoring of patients. Individual ML algorithms perform differently on different datasets. Due to the predictive results varying, this might impact the overall results. The variation in prediction results looms large in the clinical decision-making process. Therefore, it is essential to understand the different ML algorithms used to handle IoT data in the healthcare sector. This article highlights well-known ML algorithms for classification and prediction and demonstrates how they have been used in the healthcare sector. The aim of this paper is to present a comprehensive overview of existing ML approaches and their application in IoT medical data. In a thorough analysis, we observe that different ML prediction algorithms have various shortcomings. Depending on the type of IoT dataset, we need to choose an optimal method to predict critical healthcare data. The paper also provides some examples of IoT and machine learning to predict future healthcare system trends.

Keywords: IoT; ML; health prediction system; classification; prediction; supervised learning



Citation: Aldahiri, A.; Alrashed, B.; Hussain, W. Trends in Using IoT with Machine Learning in Health Prediction System. *Forecasting* **2021**, *3*, 181–206. <https://doi.org/10.3390/forecast3010012>

Received: 18 January 2021

Accepted: 3 March 2021

Published: 7 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. 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

Health prediction systems help hospitals promptly reassign outpatients to less congested treatment facilities. They raise the number of patients who receive actual medical attention. A health prediction system addresses the common issue of sudden changes in patient flows in hospitals. The demand for healthcare services in many hospitals is driven by emergency events like ambulance arrivals during natural disasters and motor vehicle accidents, and regular outpatient demand [1]. Hospitals missing real-time data on patient flow often strain to meet demand, while nearby facilities might have fewer patients. The Internet of Things (IoT) creates a connection between virtual computers and physical things to facilitate communication. It enables the immediate gathering of information through innovative microprocessor chips.

It is worth noting that healthcare is the advancement and preservation of health through the diagnosis and prevention of disorders. Anomalies or ruptures occurring below the skin periphery can be analyzed through diagnostic devices such as SPECT, PET, MRI, and CT. Likewise, particular anomalous conditions such as epilepsy and heart attack can be monitored [2]. The surge in population and the erratic spread of chronic conditions has strained modern healthcare facilities. The overall demand for medical resources, including nurses, doctors, and hospital beds, is high [3]. In consequence, there is a need to decrease the pressure on healthcare schemes while preserving the quality and standards

of healthcare facilities [4]. The IoT presents possible measures to decrease the strain exerted on healthcare systems. For instance, RFID systems are used in medical facilities to decrease medical expenses and elevate healthcare provision. Notably, the cardiac impulses of patients are easily monitored by doctors via healthcare monitoring schemes, thus aiding doctors in offering an appropriate diagnosis [5]. In a bid to offer steady transmission of wireless data, various wearable appliances have been developed. Despite the advantages of the IoT in healthcare, both IT experts and medical professionals worry about data security [6]. Consequently, numerous studies have assessed the integration of IoT with machine learning (ML) for supervising patients with medical disorders as a measure of safeguarding data integrity.

The IoT has opened up a new era for the healthcare sector that enables professionals to connect with patients proactively. The IoT with machine learning evaluates emergency care demands to make a strategy to deal with the situation during specific seasons. Many outpatient departments face the problem of overcrowding in their waiting rooms [7]. The patients who visit hospitals suffer from varying conditions, with some requiring emergency medical attention. The situation is further exacerbated when patients with emergency care needs have to wait for a lengthy queue. The problem is aggravated in developing countries with under-staffed hospitals. Many patients commonly return home without receiving medical treatment due to overcrowding at hospitals.

Yuvaraj and SriPreetha created a wearable medical sensor (WMS) platform made up of different applications and utilities [8]. The authors comprehensively analyzed the application of WMSs and their advances and compared their performance with other platforms. The authors discussed the advantages brought about by the applications of these devices in monitoring the health of patients with conditions such as cardiac arrest and Alzheimer's disease. Miotto et al. proposed a monitoring system that relies on a wireless sensor network (WSN) and fuzzy logic network [9]. Specifically, the researchers integrated micro-electro-mechanical systems (MEMS) set up with WSN to create a body sensor network (BSN) that regularly monitors abnormal changes in patients' health. Notably, the authors developed a clinical data measuring system using devices such as a microcontroller, pulse, and temperature sensor [10]. Additionally, the proposed system was integrated with base station appliances to remotely regulate the pulse and temperature of patients as well as convey the patient's data to the medical practitioner's phone. Notably, the system can send an SMS to both the patient's relatives and medical experts in emergency scenarios [3]. Therefore, the patients can acquire a remote prescription from medical practitioners using this system.

Moreover, the IoT application has made it possible for hospitals to monitor the vital signs of patients with chronic conditions [11,12]. The system uses such information to predict patient health status in different ways. IoT sensors are placed on the patient's body to detect and recognise their activity and to predict the likely health condition. For example, the IoT sensors system monitors diabetes patients to predict disease trends and any abnormal status in patients. Through the health prediction system, patients can receive suggestions of alternative hospitals where they might seek treatment. Those who do not want to visit other facilities can choose to stay in the same facility but face the possibility of long waiting queues or returning home without treatment. Rajkomar et al. [13] proposed a Zigbee Technology-hinged and BSN healthcare surveillance platform to remotely monitor patients via clinical sensor data. In particular, they utilized standards such as Zigbee IEEE 802.15.4 protocol, temperature signals, spirometer data, heart rate, and electrocardiogram to assess the health status of patients [14]. The acquired data are then relayed via radio frequencies and displayed on visual appliances including desktop computers or mobile devices. Therefore, the proposed platform could monitor attributes of patients including temperature, glucose, respiratory, EEG (electroencephalogram), ECG (electrocardiogram), and BP (blood pressure), and relay them to a database via Wi-Fi or GPRS. Once the sensor data are offered to the Zigbee, they are conveyed to a different network, permitting their visualization on appliances such as emergency devices and the mobile phones of doctors and relatives [10]. Accordingly, the integration of IoT with machine learning eases the

management of healthcare in patients by enhancing the connection between patients and doctors.

The IoT offers systems for supervising and monitoring patients via sensor networks made up of both software and hardware. The latter includes appliances such as the Raspberry Pi board, blood pressure sensors, temperature sensors, and heart rate sensors. The software process entails the recording of sensor data, data cloud storage, and the evaluation of information stored in the cloud to assess for health anomalies [15]. Nonetheless, anomalies usually develop when there exist anonymous activities in unknown body parts. For instance, the heartbeat tends to be elevated when seizures occur in the brain [16]. As a result, machine learning techniques are applied to integrate the heart rate sensor with Raspberry Pi boards to display abnormal results via either an LCD or a serial monitor. Due to the vast volume of data, cloud computing is applied to store the information and enhance data analysis [17]. Various open-source cloud computing platforms are compatible with the Raspbian Jessie and Raspberry Pi board [18]. These devices utilize machine learning algorithms to assess the stored data to recognize the existence of any anomalies [19]. Therefore, the application of machine learning in IoT helps in predicting anomalies resulting from unrecognized activities in different body parts.

It is paramount to note that machine learning is an artificial intelligence (AI) discipline. The primary objective of machine learning is to learn from experience and paradigms. In contrast to classical techniques of simply generating code, big data are input to the generic algorithm and analysis conducted using available data [20]. Big data allow the IoT and machine learning systems to easily train a system by applying simple data for predicting medical anomalies. The accuracy of predictions is directly proportional to the quantity of big data trained [21]. Therefore, big data enhance the prediction ability of machine learning techniques utilized in healthcare prediction platforms.

Fortunately, patient load prediction models are based on machine learning for prompt patient load information sharing among hospitals. In a hospital, the historical data are captured and used to forecast the future patient load to ensure adequate preparation. IoT devices with embedded machine learning methods are used to train a classifier that can detect specific health events such as falls among elderly patients. The clustering algorithms can effectively identify abnormal patterns of behaviour among patients and send out alarms to healthcare providers. Similarly, the daily activity of a patient is monitored through daily habit modelling with IoT microchips. The information is utilised for detecting anomalies among older adults.

This paper intends to analyse the most well-known ML algorithms for the classification and prediction of IoT data in the healthcare sector. We have analysed their working while comparing them based on different parameters. The study further compares existing literature, highlights their features and shortcomings, and discusses possible gaps in each approach in order to select appropriate algorithms for building an efficient prediction model. From this research, we find that K-Nearest Neighbor (KNN) may be the most popular algorithm for classification and prediction task. However, it could take a long time to predict the output in real-time applications. Therefore, some researchers have claimed that combining Long Short-Term Memory Neural Network (LSTM) with recurrent neural networks (RNN) might improve the prediction performance. In this research we are addressing the following question: How can IoT data with machine-based algorithms develop a better healthcare prediction system?

The rest of the paper is organized as follows: Section 2 discuss the ML models and classification. Section 3 discusses the most recognisable ML algorithms that are used for variety and prediction application. Section 4 discusses ML algorithm applications. Section 5 describes the use of the IoT and ML in the healthcare sector. Finally, Section 6 concludes the paper with further research directions.

2. ML Algorithms and Classification

Machine Learning (ML) is a phenomenon associated with the field of AI. ML equips a system with the ability to automatically analyse and understand a combination of inputs as an experience without the need for any additional help [22]. There are two critical phases of building an efficient ML model: training and testing. The training phase (a highly research-intensive phase) involves providing labelled or unlabelled inputs to the system. The system then stores these training inputs in the feature space to refer to for future predictions. Finally, in the testing phase, the system is fed an unlabelled input for which it must predict the correct output.

Simply put, ML uses known data in its feature space to predict outcomes for unlabelled data. Hence, a successful ML model can refer to its experiences and understandings to predict outputs. The accuracy of such a model depends on the accuracy of its output as well as on model training.

In recent times, machine learning programs are widely applied in healthcare service applications. Such machine learning algorithms are also applied in many clinical decision support systems to establish advanced learning models to enhance healthcare service applications. Support vector machines (SVM) and artificial neural networks are examples of the integration of ML in healthcare service applications. These models are used in various cancer classification applications for the accurate diagnosis of cancer type. These algorithms work by evaluating the data obtained from sensor devices and other data sources. These algorithms identify the behavioural patterns and clinical conditions of a patient. For example, these algorithms identify improvements in a patient, habits, variations in daily routine, changes in patterns of sleeping, eating, drinking, and digestion, and changes in patient mobility. The behavioural patterns determined through these algorithms can then be used by healthcare applications and clinical decision support systems to suggest changes in patient lifestyle and routines and recommend various specialised treatments and healthcare plans for patients. This enables doctors to develop a care plan to ensure that patients introduce the recommended changes in their lives.

ML technology has three main model types: supervised learning, semi-supervised learning, and unsupervised learning. Each ML type has several common algorithms, as Figure 1 shows. This section introduces the most popular ML methods employed for prediction and classification purposes. These methods are K-Nearest Neighbor, Naïve Bayes, Decision Trees, Support Vector Machine (SVM), Neural Networks, Gradient Boosted Regression Tree, and Random Forest in Supervised Learning. All these methods will be discussed in Section 3. However, before discussing these methods, the paper will first introduce the idea of data points and data labels in the context of machine learning.

2.1. Data in ML

Figure 2 clearly shows that data and datasets must be collected from many sources for developing an analytic model using ML technology. The datasets acquired are saved in a conventional, centralised way in the cloud. Data points, also referred to as samples or observations, are the basic units of the dataset. These data points are representative of a system unit. This system unit is evaluated to construct the training datasets. A data point can indicate a patient's information regarding a cancer tissue sample or any other thing.

Recently, there has been a huge increase in the accessibility of data points from healthcare institutions. Such data points can be labelled or unlabelled. Labelled data have a distinctive feature assigned to them (which is called a label), which can also be referred to as an output or response. It is also referred to as a dependent variable as far as the classical statistical literature is concerned. A label can be either categorical or ordinal, where the categorical one has no set order of predefined values (e.g., male and female). In contrast, the ordinal has a basic order of predefined values (e.g., disease stage). Moreover, a label can be a numerical value such as real numbers.

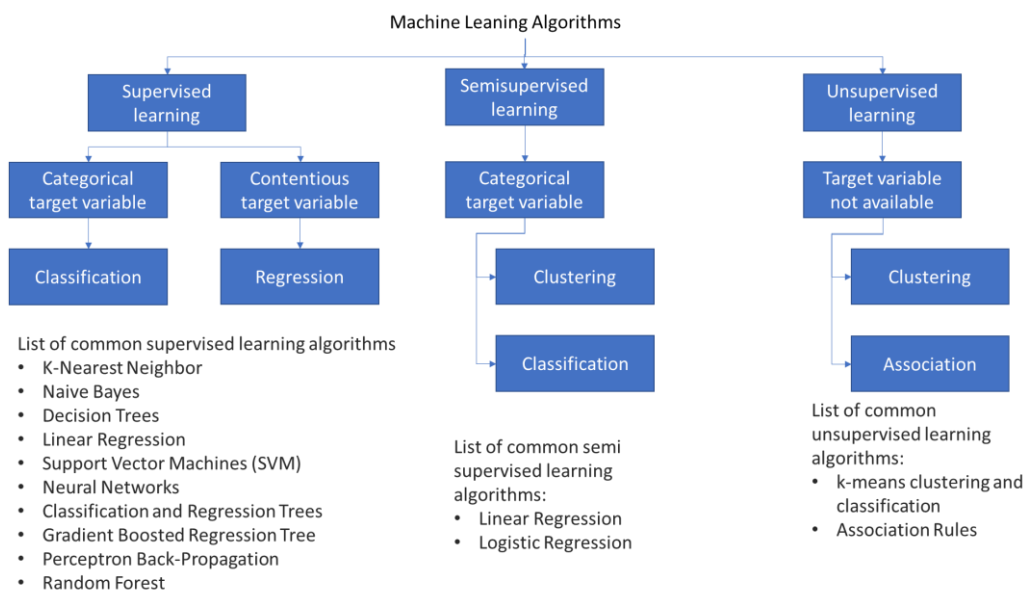


Figure 1. Machine learning algorithms’ classification.

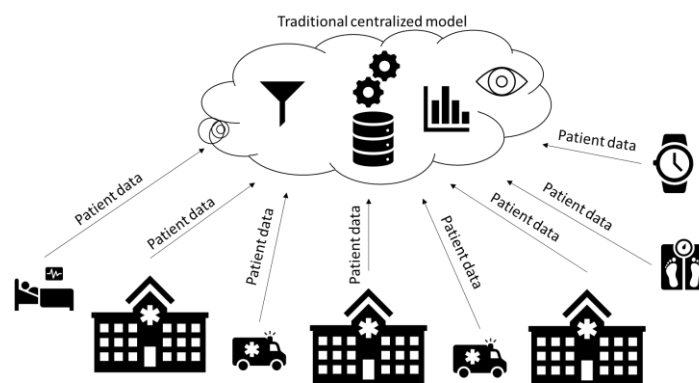


Figure 2. Traditional centralised learning: ML runs in the cloud, gathering data from different hospitals.

Although most ML models can employ both kinds of data, labelled and unlabelled, in specific conditions, labelled data are used in supervised learning. In contrast, unlabelled data are employed in unsupervised learning, while semi-supervised learning can use both labelled and unlabelled data. In the next section, we will attempt to further elaborate on each data type and their use in every ML model. Table 1 presents a list of alternative words used in the paper to indicate the same meaning.

Table 1. Terminology used in the study.

Terminology	Alternative Word
Datapoint	Input, observation, and sample
Label	Output, response, feature, and dependent variable

2.2. Machine Learning Algorithm Classification

The following three basic machine learning models will be analysed in the paper: supervised learning, semi-supervised learning, and unsupervised learning.

2.2.1. Supervised Learning

The most important ML model is the supervised learning model. It is mainly used to cater to real-world applications [23]. This model is used to predict outcomes from certain

sets of given input and a pair of input/output examples. A pair of input objectives, an input vector, and the desired output value called a supervisory signal are all involved in each supervised training dataset. These examples are used to train ML algorithms to obtain an inferred classifier function after analysing the training datasets. The objective of training algorithms in supervised learning is to predict the value of one or many outcomes through various input features.

A distinguishing feature of the supervised learning model is human involvement in it. Human involvement is essential in the beginning to construct a dataset, which later works on its own by generalising and learning from examples fed through input. For the construction of a dataset, first pairs of inputs and preferred outputs are provided to the ML model. This model then finds a way to work independently to generate outputs. The main problem arises when the model has to predict the output for a new input independently without human assistance. Hence, ensuring the accuracy of the proposed model is essential.

Despite the evident effectiveness of supervised learning, it has the drawback that it requires numerous labelled data to develop a large-scale labelled dataset [24]. Supervised learning models are used extensively for classification and regression purposes. This paper, however, will merely discuss the classification approach.

Classification or prediction is the main objective of the use of machine learning methods. These methods classify and foretell class labels by making use of a preset list of examples. Classification samples either completely belong to a certain class or do not belong to a class at all. They do not belong to any class partially. Missing values negatively affect classification and prediction processes.

There are two forms of classification: binary and multiclass. Binary classification deals with two sets of classes, and the input data are arranged in these two sets of classes. An example is making a YES/NO prediction or classifying e-mails into two categories, spam and no spam. These classification classes are interpreted as 0 and 1. Conversely, multiclass classification is concerned with three or more predictable classes. An example would be the identification of cancer stage. Here, classes are defined as 0, 1, 2, etc.

Figure 3 shows how to solve a given problem using supervised learning. There are usually particular steps that need to be followed. First, the type of training example should be determined. Next, the training set needs to be gathered, either from human experts or from measurements. It also needs to be representative of the real-world use of the function. Then, the representation of the input feature must be determined. The representation should contain enough information to accurately predict the output. This is followed by selecting the learning algorithm. After completing the design, we run the learning algorithm on the gathered training set. At this stage, some algorithms in supervised learning are required from the user to determine certain control parameters, especially for prediction problems [25]. These parameters are called a validation set, which may be adjusted by optimizing performance on a subset of the training set, or via cross-validation [26]. Cross-validation has two types: exhaustive and nonexhaustive. In exhaustive cross-validation, the methods learn and test all possible ways to divide the original sample into training and validation sets. Examples of exhaustive cross-validation methods are leave-p-out cross-validation and leave-one-out cross-validation. In contrast, nonexhaustive cross-validation methods do not compute all ways of splitting the original sample. Those methods are approximations of leave-p-out cross-validation. k-fold cross-validation, 2-fold cross-validation, and repeated random subsampling validation are examples of such methods. Finally, the accuracy of the learned function should be evaluated. After parameter adjustment and learning, the user uses a test set that is separate from the training set to measure the function result performance.

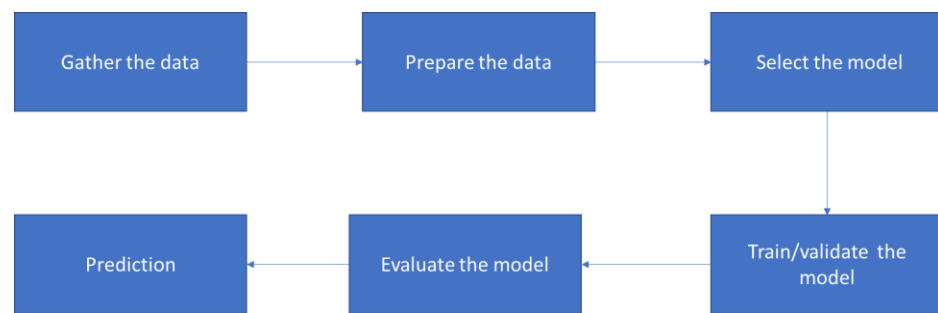


Figure 3. Steps to solve a problem using supervised learning.

2.2.2. Unsupervised Learning

One of the functions of unsupervised machine learning is identifying hidden structures within data that have not been marked. Several successful applications have utilised this; however, these applications are often difficult to evaluate. This is due to a lack of training regarding the use of unsupervised machine learning. As a result, there is a lack of error or reward indicators for analysing prospective solutions. Here, the reward signal serves as a distinguishing factor for supervised and unsupervised ML. In the field of statistics, unsupervised learning is used for density approximation. The neural network (NN) models [22], the self-organising map (SOM), [27], and adaptive resonance theory (ART) [28] also make use of unsupervised learning.

Unsupervised learning includes the transformation of datasets and clustering. In the transformation process, data in the dataset are altered to present them in a different, new form so that they become easy to understand for humans and machine algorithms. Clustering algorithms, on the other hand, separate datasets into significant groups of related objects. The K-means clustering is the most well-known and simplest unsupervised algorithm, and identifies clusters of similar data. There are two steps to this algorithm: the first step is allocating each data point to the nearest cluster centre, while the second step is fixing all cluster centres as the mean of data points that are allotted to them.

One main problem in unsupervised learning is evaluating its success. The success of unsupervised learning tells if the algorithm has learned useful things or not. Labels or outputs are not provided in unsupervised learning; hence, the right output is not known. Hence, it becomes difficult to determine the performance of the algorithms. This is why unsupervised learning is used solely in an exploratory way, e.g., for better comprehension of data. Another critical feature of unsupervised algorithms is the preprocessing step for supervised algorithms. Finding a new form of data representation can improve the accuracy of supervised algorithms.

2.2.3. Semisupervised Learning

One of the ML model branches is the learning technique, which depends on marked and unmarked data to equip the ML model. In a real-life scenario, minor marked data must be used with a huge amount of unmarked data to obtain greater learning accuracy. The tagging of a dataset requires human involvement. The tagging procedure is time-consuming, which might hinder creating completely labelled training and bring about heavy expenses. As a result, in certain instances, semisupervised learning may prove to be a better solution.

In the case of a limited number of labelled samples, semisupervised learning is mostly used to enhance the model's performance. Currently, there are numerous unlabelled samples available. These unlabelled samples can be utilised to enhance the performance of the model. The poor model performance is more obvious than the improvement and is caused due to the implementation of the unlabelled sample data in semisupervised learning. As a result, semisupervised learning is not widely used in applications; supervised learning, which shows top performance in machine learning problems, is preferred [29].

The distinctions between the supervised, unsupervised, and semisupervised learning models are outlined in Table 2.

Table 2. The difference between supervised learning, unsupervised learning, and semisupervised learning.

Learning Class	Data Type	Usage Type	Output Accuracy/ Performance	Affected by Missing Data	Scalable	Cost
Supervised	Labelled	Classification Regression	High	Yes	Yes, but we need to label large volumes of data automatically.	Expensive
Unsupervised	Unlabelled	Clustering Transformations	Low	No	Yes, but we need to verify the accuracy of the predicted output.	Inexpensive
Semi-Supervised	Both Labelled and unlabelled	Classification Clustering	Moderate	No	It is not recommended.	Moderately priced

3. Commonly Used Machine Learning Methods

For classification and prediction, researchers have developed and adopted several famous machine learning models. Some of them are explained in the sections below.

3.1. K-Nearest Neighbor (K-NN)

The basic machine learning model that is popular for classifications and regression tasks is K-Nearest Neighbor (K-NN). This model's main focus is determining the distance between a new unlabelled data point and the existing training datasets stored in the feature space. It is essential for the prediction of the class [30]. In this process, the nearest data points will be ordered according to the k-value of the new observation. The k-value is a hyperparameter of the following model, which is also utilised to sort the new observation's k-nearest data points. The K-NN classifier votes and allocates the predicted class to the new unlabelled data sample depending upon the class label's volume in k-neighbours. In an optimised K-NN algorithm, a neighbour has only positive relationships with the requestor. Such algorithms are commonly used in multiple types of research [31,32].

In the mentioned method, it is essential to evaluate the weight of influence of the adjacent neighbour. The nearest neighbour can affect more than a distant neighbour. The K-NN algorithm uses the distance function to determine the weight of influence of the adjacent neighbours. Euclidean Distance, Manhattan Distance, Pearson Correlation, and Spearman Correlation are typical distance functions used for continuous variables. However, the Hamming distance function is commonly utilised to evaluate the number of disparities in attributes such as the characteristics of the two data points. Figure 4 illustrates how the K-NN classifier predicts the class label for the new data point by using the Euclidean distance function. The k-value in this example is six data points, so the new sample's predicted class is the most repeated class within these six data points, which means that the new data point sample might belong to the blue class in this example.

There are pros and cons to using the K-NN algorithm. A nonparametric approach is one of the advantages of using the K-NN classifier, which implies no hypothesis for the fundamental distribution of data. For instance, the structure of the model is established upon the dataset. Another advantage is that it is easy to understand and easy to incorporate. It can update its set of labelled observations to modify because K-NN is not explicitly trained quickly.

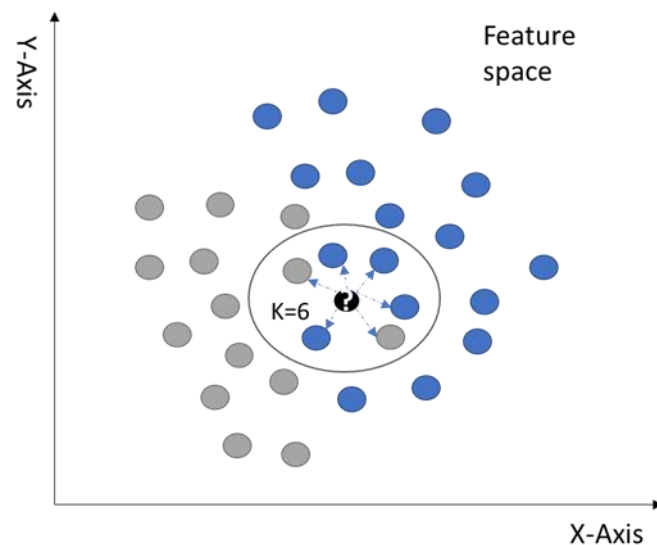


Figure 4. Classification scenario using a K-NN algorithm.

One of the cons is that the K-NN model requires a large amount of time to execute due to each new observation's comparison process with the labelled observation. This results in inferior performance if it is used on disproportionate datasets. Additionally, K-NN is highly dependent on its k hyperparameter. It will produce highly accurate results on the training set using only a single neighbour for classification, but the problem is that the model is very complicated and there is a high probability of computational complexity.

Thus, it is important to accurately choose the number of neighbours for the purpose of classification, so that the K-NN model is neither overcomplicated nor oversimplified and can be easily generalised. It is also necessary to apply homogenous features rather than heterogeneous features. When the model uses the K-NN algorithm, the distance between the samples with the heterogeneous features can be easily affected. This is due to the size and changes in certain features, which causes loss of information from other components. Consider the following examples. There are three features, x_1 , x_2 , and x_3 , of which x_1 and x_2 are normalised variables having values from 0 to 1. In comparison, the value of x_3 lies in the range of 1-100. The effect of x_1 and x_2 on any distance function, like the Euclidean distance function, will be marginal when evaluating the distance between the two data points.

On the other hand, x_3 will be higher than the evaluated distance values. The value of x_3 is kept within the range of 0 and 1, which further highlights the evaluated distance values. Likewise, it is possible to customise and implement the distance function, which is not highly dependent upon heterogeneous features. In summary, it is easy and less time-consuming to develop the K-NN model, but the set predictions can be very slow in training if there are many features or samples involved in the learning process. There are many real-world implementations of the K-NN algorithm [33], such as gene expression, image recognition, video recognition, and pattern recognition.

3.2. Naïve Bayes Classification (NBC)

The naive Bayes classifier (NBC) is a basic classifier used for probabilistic classification; it has been developed based on the Bayes theorem. The NB model is based on the assumption that each feature is statistically independent of the other features and irrelevant to them in the training set [34]. These assumptions are used to predict the class of new observations through equations. The NB classifier evaluates the probability of a new unlabelled sample B ($B = b_1, b_2, b_3, \text{etc.}$) of becoming part of class A . The output of the model is predicted on the basis of the probability with the greatest $P(A|B)$. $P(B)$'s value does not impact the selection of the class having the greatest $P(A|B)$. Moreover, the relative

frequency of class A helps to determine the value of $P(A)$ with the help of the data points given in the datasets used for training.

$$P(A|B) = \frac{P(B|A) * P(A)}{P(B)} \quad (1)$$

Supervised learning is used in the NB classifiers to enhance the efficiency of learning classification tasks. This higher efficiency is because NBC can be trained to comprehend the parameters by extracting basic per-class statistics from features simply by performing individual observation of these features. The NB model utilises the parameter approximation tool through the maximum likelihood approach in various applications. Hence the functioning of the NB model may be independent of the Bayesian algorithm.

There are several real-world uses of the NBC model, including real-time predictions as well as text classifications. The NB classifier also has some major drawbacks, although they are characterised by simple structure and presumptions. One of these drawbacks is that they only need a limited amount of data to estimate the necessary categorisation parameters. Moreover, there is no need to evaluate the entire covariance matrix, but it is important to evaluate each class's variance.

3.3. Decision Tree (DT) Classification

The simple classification algorithm consisting of the internal node and one class-labelled leaf node is called a DT (Decision Tree) [35]. The solution of classification issues in the DT approach is done through constant splitting of the input space to create a tree with pure and straightforward nodes and points related to a single class. As we move down this tree, a new point is classified by selecting a single (side) branch of the tree at each point. The development of decision trees is dependent on the kind of target variable in the current model. The DTs algorithm utilises the reduction in variance approach to form a tree model with continuous variable test points. In contrast, the Gini impurity approach is used for categorical target variables.

The DT technique initiates at the root node and terminates at the last node, known as the terminal or leaf node of the tree. This entire method is shown in Figure 5. The internal nodes are present amid the root and leaf nodes and are used to test the data point characteristics. There is one potential outcome for each internal node. Consider the following example. There is an unlabelled new data point that begins with the root node. It approaches the next node based on the outcome of the root test. The same strategy is followed and the subsequent node is an internal node. After that, the new unlabelled data point is allocated to the leaf node and the prediction of a class of the new unlabelled data point is done based on a class label corresponding to the concerned leaf node.

There are several real-world uses of DTs in various sectors. In the healthcare sector, DTs are used for early diagnosis of cognitive disability, which increases the efficiency of screening positive cases; it also determines key risk factors for the potential occurrence of various dementia types. The Sophia robot developed in Saudi Arabia is one of the DT algorithm applications devised to interact with humans. It is also a well-known algorithm in machine learning.

The DT model has benefits and drawbacks, just like the other ML models. The advantages offered by the DT algorithm are that it is simple to interpret. It can also deal with continuous and categorical attributes and needs little to no information processing. Moreover, this method is significant when it comes to classification tasks and making predictions. However, in the case of misbalanced datasets, the DT can have an overall negative performance. This model is also noise-sensitive and can also lead to overfitting if there is noise in the training dataset.

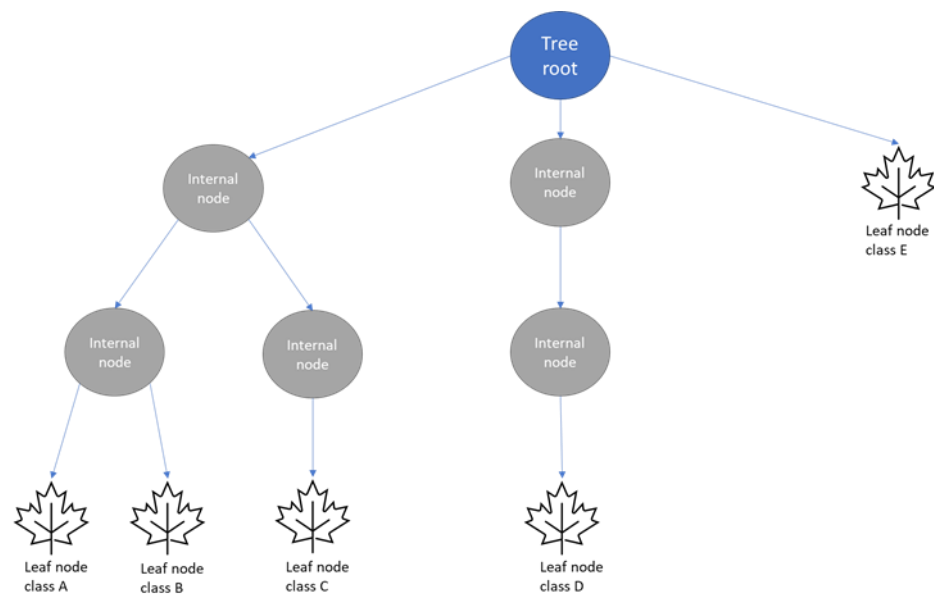


Figure 5. Decision trees (DTs).

Furthermore, any rise in dataset inputs will lead to greater model complexity. Hence, huge memory is essential for large datasets. This challenge can be overcome by restricting the height or depth of the tree or fixing the sample quantity essential for additional classification. Low-intensity decision trees are valuable; however, they may lead to certain drawbacks for the model, including low variance and high partiality, while the opposite is true for high-depth decision trees. Bagging and boosting in decision trees are common methods to minimise high model variance or high bias.

3.4. Random Forest

One of the ensemble models is a random forest that integrates multiple models and shows compatibility with an extensive range of datasets for classification or regression [36]. Bootstrap aggregation or bagging are examples of such models. Bagging can reduce model variance and consequently enhance generalisation to prevent overfitting. This model can overcome high conflict even beyond the decision trees.

There are several decision trees (DTs) present in a random forest model and each of them shows slight disparity from the others. For a specific data point, various outcomes obtained from each decision tree will be integrated. In the integration process, a majority vote is obtained in the case of classification, and the average value is obtained when it comes to regression tasks. The performance of combined decision trees is better than a single decision tree in terms of predictions because of the separate training of all the decision trees on random samples taken from a training dataset. This model is known as random forest since it involves randomisation in tree building to make certain that all trees are different from each other. There are two different ways to randomly produce a training set for randomisation of the trees in a random forest: The first method is to choose attributes from each test. The second method involves a selection of data points that were involved in the development of a tree. This will be followed by the training of a decision tree by using the developed training set and the chosen attributes. In general, applying such a strategy lowers relationships within the decision trees and enhances the performance of the model.

3.5. Gradient-Boosted Decision Trees

Another form of ensemble model is the gradient-boosted decision tree [37]. Gradient boosting is similar to a random forest model since both are strong models involving multiple decision trees. Both classification and regression task may be performed through

this model. However, this model is different from the random forest model in that it lacks randomisation in the process of developing model trees. Prepruning is used in this model, i.e., trees are constructed serially, and every tree attempts to rectify the mistake associated with the preceding tree. In the gradient-boosted model, the depth of trees is quite small, with values of depth ranging from 1 to 5; consequently, we get smaller memory and faster predictions. The basic concept of gradient boosting is to mix several easy models (or weak learners). For instance, in a shallow tree, every tree can only make effective predictions for its corresponding data, suggesting that a higher number of trees will lead to better overall performance. The most extensive application of gradient-boosted decision trees can be seen in supervised learning due to their robust nature. Their main disadvantage is that they need cautious standardisation of the parameters and involve more training time. Moreover, they fail to be effective when data points are located in high-dimensional space.

3.6. Support Vector Machines (SVMs)

Cortes and Vapnik [38] presented the support vector machine (SVM), commonly referred to as a kernelized support vector machine (KSVM). SVMs can be defined as a supervised machine learning methodology that develops the decision boundary (or hyperplane) among multiple classes by performing analysis of various inputs within a dataset; hence, it becomes possible to predict labels on the basis of single or multiple feature vectors. Its position is such as to ensure the greatest possible distance from data points close to each class. The name support vector machine has been derived from these closest points, which are also known as support vectors.

The main objective of this technique is binary linear categorisation and prediction. Numerous biological applications have made effective use of this technique [39]. SVMs are most commonly used in biomedical practice to automatically categorise profiles for microarray gene expression. Moreover, substances with high diversity like protein and DNA sequences, as well as mass spectra and microarray gene expression profiles, can be classified with the help of SVMs. The kernel method is another important application of SVMs, which helps to model higher-dimensional and nonlinear models. In the case of a nonlinear model, such as in Figure 6D, nonlinear prediction and classification can be done through the kernel technique based on SVMs, which helps with the mapping or plotting of cases within high-dimensional space. Calculations in high-dimensional space may be time-consuming; however, the use of a kernel function results in quick calculations.

In the process of mapping, the SVM indicates the decision boundary between two classes by determining the value of each data point, particularly the support vector points located at the boundary between the concerned classes. The SVM learns the significance of every data point in the course of mapping; in particular, the significance of support vector points located at the border among the classes is evident for SVMs to indicate the decision boundary among those classes. The distance to every support vector is measured for the prediction of a new point. KSVMs perform classification by considering the distance to support vectors and the significance of support vectors. It is imperative to note that significant improvement may be brought about in the performance of KSVMs by data scaling in the range of 0 to 1.

Figure 6A–C illustrate a case of classification task whereby a simple SVM algorithm is employed; we assume that there are numerous cases in a dataset. Moreover, it is presumed that single or multiple classes are associated with each case. The cases within the given dataset are spread into a couple of classes by the SVMs by developing a model that selects the most appropriate hyperplane linearly. In the context of SVMs, each case is shown by a point in space such that the point plots the cases in two distinct classes with a gap between them. Moreover, the new case's location with reference to the gap is used to plot these cases into space and helps with their classification into single or multiple classes.

The middle line in the gap represents the SVM decision surface. Any data point located above the decision surface will be classified as class 1 (square); otherwise, it will belong to the other class (star). Sometimes, real datasets cannot be separated easily when

the dataset contains an error. Therefore, SVM can correct errors in the data if the SVM algorithm has to be modified by adding a soft margin. Consequently, some data points push their way through the separating hyperplane margin without affecting the final result.

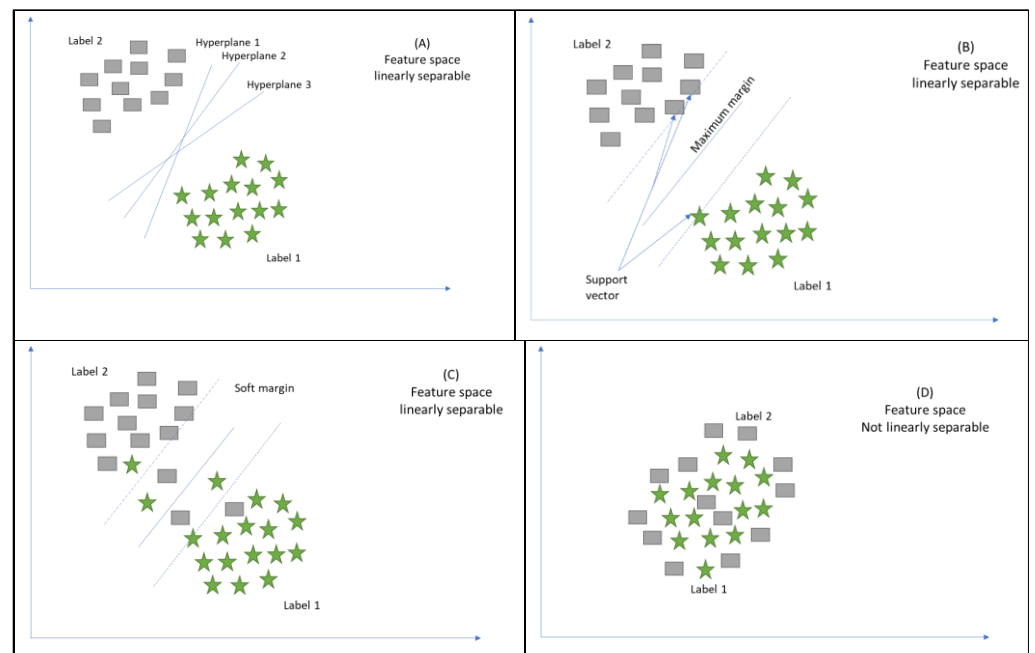


Figure 6. Classification scenario using a simple SVM algorithm.

3.7. Neural Networks

Nowadays, the availability of abundant data and the power of computing, along with advances and innovations in deep learning, have influenced the success of the current ML model. The reason behind this success is that neural networks contain a large number of computational units that can map the given input explicitly with the predicted output in the case of using nonlinear functions. Therefore, neural networks are better than other machine learning algorithms if enough time, data, and careful tuning of the parameters can be ensured. The most widely used deep learning approaches are deep multilayer perceptron (MLP), recurrent neural networks (RNN), and convolutional neural networks (CNN). MLP is the most popular neural network. It is a powerful modelling tool that uses examples of data with known outputs like supervised learning. It also generates a nonlinear function model to predict the output from the given input [40,41].

In contrast, the CNN approach involves numerous layers, including a convolutional layer, nonlinearity layer, pooling layer, and fully connected layer. Both convolutional and fully connected layers have parameters, while the pooling and non-linearity layers do not have parameters. The performance of such an approach is exceptional, especially in machine learning problems that deal with image classification dataset (Image Net) [42], computer vision, and natural language processing (NLP) [43].

Finally, RNN is by nature recurrent because it performs the same procedure for every data input, while the predicted output of the current input will be learned from the previous computation. After detecting the output, the result will be copied and sent back into the recurrent network. However, this model has some limitations. Therefore, a Long Short-Term Memory Neural Network (LSTM) has been combined with RNN in order to overcome the instability of long-term predictions because of either an exploding or a vanishing gradient [44]. Reviewing these approaches in detail is beyond the scope of this review.

All the reviewed ML algorithms have been summarized in Table 3.

Table 3. Summary of the reviewed ML algorithms.

Algorithm Name	Learning Type	Used for	Commonly Used Method	Positives	Negatives
K-Nearest Neighbor (K-NN)	Supervised	Classification, Regression	Continuous variables (Euclidean distance) Categorical variables (Hamming distance)	Nonparametric approach. Intuitive to understand. Easy to implement. Does not require explicit training. Can be easily adapted to changes simply by updating its set of labelled observations.	Takes a long time to calculate the similarity between the datasets. The performance is degraded because of imbalanced datasets. The performance is sensitive to the choice of hyperparameter (K value). The information might be lost, so we need to use homogeneous features.
Naïve Bayes (NB)	Supervised	Probabilistic classification	Continuous variables (Maximum likelihood)	Scanning of data by looking at each feature individually. Collecting simple per-class statistics from each feature helps with increasing the assumptions' accuracy.	Requires only a small amount of training data. Determines only the variances of the variables for each class.
Decision Trees (DTs)	Supervised	Prediction, Classification	Continuous Target Variable (Reduction in Variance) Categorical Target Variable (Gini Impurity)	Easy to implement. Can handle categorical and continuous attributes. Requires little to no data preprocessing.	Sensitive to the imbalanced dataset and noise in the training dataset. Expensive, and needs more memory. Must select the depth of the node carefully to avoid variance and bias.
Random Forest	Supervised	Classification, Regression	Bagging	Lower correlations across the decision trees. Improves the DT's performance.	Does not work well on high-dimensional, sparse data.
Gradient Boosted Decision Trees	Supervised	Classification, Regression	Strong prepruning	Improves the prediction performance iteratively.	Requires careful tuning of the parameters and may take a long time to train. Does not work well on high-dimensional, sparse data.
Support Vector Machine (SVM)	Supervised	Binary classification, Nonlinear classification	Decision boundary, Soft margin, Kernel trick	More effective in high-dimensional space. Using the kernel trick is the real strength of SVM.	Selecting the best hyperplane and kernel trick is not easy.

4. Machine Learning Applications

Identifying the right problem to be solved using ML is the first step in building a machine learning product. Healthcare is a data-rich environment. Even though a model can be created to produce understanding, it must have the potential to impact patient care. In this section some of the applications that use the productive model are listed below. At the end of this section, Table 4 summarizes the reviewed applications.

4.1. Medical Imaging

Machine learning finds applications in medical imaging, which refers to the processes and methods utilized to create images of body parts for treatment and diagnostic purposes. Some of the imaging techniques implemented today include magnetic resonance imaging (MRI) and X-ray radiology. The current practice entails taking these images and having a health professional examine them manually to determine abnormalities. This process is not only time-consuming but also prone to errors. Accordingly, the utilization of machine learning algorithms improves the accuracy and timeliness of disease prediction, detection, and diagnosis [45]. Researchers have demonstrated how various machine learning algorithms such as artificial neural networks (ANN) can be integrated into medical imaging to enable computer-aided disease prediction, diagnosis, and detection [45]. Deep learning approaches, particularly convolutional neural networks (CNNs), have emerged as effective tools for image and video analysis, which is central to medical imaging [11]. The input data type for medical imaging applications is mostly images, such as X-rays and CT scans [45,46]. The IoT devices used in a machine learning setup include X-ray machines and CT scanners, which are readily available in healthcare settings [46]. The application of machine learning to medical imaging largely adopts the supervised learning approach.

4.2. Diagnosis of Disease

Disease diagnosis forms a critical component of care delivery as it determines the kind of intervention that should be attempted. Machine learning is applicable to disease diagnosis as it enables the examination of environmental and physiological factors for diagnosing diseases effectively. It allows for the creation of models for relating variables to a disease. In other words, machine learning can be used to identify the risk factors associated with a given disease, as well as the signs and symptoms, to improve diagnosis efficiency and accuracy. Some of the diseases currently being diagnosed using machine learning include glaucoma, age-related macular degeneration, etc. [47]. Some of the ML techniques used in diagnosing diseases include support vector machines, deep learning systems, convolutional neural networks, and backpropagation networks [47]. The input data type differs based on the disease being diagnosed. In most imaging diagnosis machine learning experiments, image data are routinely used. Similarly, time series data, including components such as demographics, gene expression, symptoms, and patient monitoring, are used for chronic disease diagnosis [48]. The application of ML can adopt either supervised or unsupervised learning approaches to learn patterns from data to enable disease diagnosis. The IoT devices and sensors utilized depend on the input data required. For imaging-based diagnosis, scanning equipment is the main IoT device. Likewise, IoT devices can be deployed to collect data such as weight, heart rate, and blood pressure to enable disease diagnosis.

4.3. Behavioural Modification or Treatment

Behavioural modification, as the name suggests, encompasses helping a patient change undesirable behaviour. Behavioural modification is an example of a treatment often prescribed to patients whose behaviours contribute to their bad health. The application of machine learning to behavioural change is made possible through the IoT, which enables the collection of vast amounts of information concerning people. Accordingly, machine learning algorithms can be used to analyse the behaviour of individuals and recommend suitable changes. In addition to providing people with alerts and notifications to influence change, machine learning algorithms can give people self-knowledge and recommend re-

sources for behavioural change. Machine learning can also be used to evaluate behavioural change interventions to determine the most effective one for a given patient [49]. Some of the machine learning algorithms applied in behavioural modification include the Bayes network classifier, decision trees, and support vector machine (SVM) [50]. The input data for these algorithms are obtained through feature extraction, which produces tabular data [50]. Accordingly, the applicable IoT devices are the ones that collect information that can be deduced to define human behaviour, such as videos, images, and recordings.

4.4. Clinical Trial Research

Clinical trials are studies performed to examine the effectiveness and safety of behavioural, surgical, and medical interventions. As clinical trials often involve human subjects and constitute the final step of the research process, they must be conducted carefully to avoid harm to the participants. Machine learning can be used to improve the clinical trial process by enabling the acquisition of knowledge concerning the effectiveness of interventions from the assessment of publicly available clinical and biomedical datasets, information obtained from health records, and practical evidence from sensors [51]. Machine learning algorithms allow healthcare professionals to examine vast amounts of data to identify insights relating to the effectiveness and safety of a given intervention. For example, ML can be applied to clinical research trials, targeting the creation of medications for COVID-19 [52]. The initial step in the implementation of ML learning algorithms in clinical trial research involves extracting features from datasets [52]. Accordingly, the input data include images and tables relating to the clinical trial. The IoT devices implemented should be able to collect data relating to the variables in the clinical trial. The typical sensor data could include weight, heart rate, blood glucose, and blood pressure.

4.5. Smart Electronic Health Records

Electronic health records, which have replaced patient charts, provide timely access to patient information, enabling care providers to offer quality care. Machine learning offers a way of integrating intelligence into electronic health records. In other words, rather than acting as storage for patient data, electronic health records can be enhanced through machine learning to include smart functions. For example, smart electronic health records can assess patient data, recommend the most appropriate treatment, and aid in clinical decision-making. In fact, the integration of machine learning with electronic health records has been shown to improve ophthalmology [53]. Additionally, smart electronic records can evaluate vast amounts of data to quantify the quality and safety of care provided in a facility and highlight areas requiring improvement. Machine learning models that can be integrated in electronic health records include linear and logistic regression, artificial neural networks, and support vector machines [53]. The input data type can include text, images, tables, and time series. For example, time series data obtained from a patient's medical record can be used to predict postpartum depression [54]. Recurrent deep learning architectures have been shown to be accurate for predicting diseases when incorporated into electronic records [55]. The IoT sensor data that are incorporated in such ML models include weight, heart rate, blood pressure, temperature, and blood glucose. The idea is that the sensor data incorporated should be symptoms of the disease or condition under consideration.

4.6. Epidemic Outbreak Prediction

Diseases that emerge and spread quickly in a community can be devastating and difficult to manage. Consequently, stakeholders within the healthcare industry recognize the need to implement tools and strategies to predict the outbreak of epidemics and prepare for them. The availability of big data allows regulators, administrators, and healthcare workers to deploy machine learning algorithms to predict epidemics. Long short-term memory (LSTM) and deep neural network (DNN) learning models are some of the machine learning algorithms used for predicting diseases [56]. The input data that can be fed into

the ML algorithms include text, time series, numerical, and categorical data. For example, time series data can be used in a machine learning setup to predict future disease trends. When predicting diseases, some of the factors fed into machine learning algorithms include population density, hotspots, vaccination levels, clinical case classifications, and geomapping [57]. Accordingly, IoT devices that could be used include satellites and drones to capture population densities and other forms of geography-related data. Weather-related data and types of information relating to the environment and that influence the possibility of epidemics can also be collected. Furthermore, clinical data obtained at the patient level, such as temperature, blood pressure, and glucose levels, are also helpful. Overall, disease surveillance is essential as it helps with preventing epidemics and allowing stakeholders to prepare for epidemics that might occur.

4.7. Heart Disease Prediction

Heart disease is a leading cause of death in most parts of the world. Due to changing lifestyles and other risk factors, the incidence of heart disease is increasing globally. In 2016, cardiovascular diseases were responsible for 17.6 million deaths globally, a rise of 14.5% as compared to 2006 [58]. A key component of managing heart disease entails being able to predict the disease and implement the right protective and treatment strategies. Machine learning offers this capability as it allows health providers to evaluate patient data and forecast the incidence of heart disease [58]. Patients who are found to be at increased risk of heart disease can be recommended interventions to avert the disease. The input data types for heart disease prediction machine learning algorithms include images, time series, text, and tabular data. For example, tabular data can be used together with algorithms such as Naive Bayes, K-NN, SVM, decision tree, and decision tables to predict heart disease [59]. The IoT sensor data that should be fed into the system relate to the risk factors of heart disease. As such, devices that can record blood pressure, heart rate, physical activity, and weight should be incorporated.

4.8. Diagnostic and Prognostic Models for COVID-19

Machine learning can also be applied in the diagnosis and prognosis of COVID-19. The idea is to develop an algorithm that accepts the predictors of prognosis and diagnosis and provides an accurate outcome. The most reported predictors include body temperature, age, lung imaging features, and lymphocyte count [60]. Machine learning algorithms are particularly effective as they can examine many lung images of patients with COVID-19 and are able to differentiate between those affected by COVID-19 and those that are not affected. Therefore, the input data type for COVID-19 prediction models include images, tabular, text, and time series. For instance, lung images can be used with ML classifiers to diagnose COVID-19 [60]. A study demonstrated high accuracy in the prediction of COVID-19 by using eight binary features: being aged 60 and above, sex, contact with an infected person, and five initial clinical symptoms [61]. Therefore, IoT sensor devices included in this machine learning setup ought to be able to measure temperature and take images of the lungs. Besides improving the accuracy of the diagnosis and prognosis of COVID-19, machine learning algorithms are fast and efficient.

4.9. Personalized Care

Offering personalized services is central to patient-centred care. Patients require care that aligns with their needs, expectations, and beliefs. In addition to improving clinical outcomes, personalized care enhances patient satisfaction and improves the utilization of formal health services. Machine learning algorithms can play a role in enabling the provision of personalized care by allowing healthcare workers to examine each patient's data and develop personalized care plans [62]. Machine learning systems harness the power of health records and integrate disparate data sources to discover person-specific patterns of disease progression [63]. The obtained information supports clinical decision-making by allowing healthcare professionals to provide personalized care. The input data

types for enabling ML personalized care can be text, time series, and tabular data. Tabular data obtained from the patient's medical record can be used to determine the best course of treatment using appropriate ML algorithms. Similarly, IoT data that can be fed into the algorithm include blood glucose, blood pressure, heart rate, and weight.

5. IoT and Machine Learning Applications in Healthcare Systems to Predict Future Trends

As previously discussed, the IoT and machine learning AI have enhanced the health sector in that patients can wear devices like premium jackets and smart bands that are used to monitor their condition and send regular reports to a database accessed by doctors and medical practitioners [64]. The devices can monitor the vital signs and organs of a patient and send out a progress report to a specific database. The system also collects and reports pathogen presence and manifestations [65]. This is a crucial advance that helps the healthcare system deliver best practices.

The availability of smart pills, sensors, and wearable monitors in healthcare adds value to the sector. These tools help with monitoring and predicting signs and future trends in disease patterns. The essence of automating the patient and disease monitoring tasks saves time and steps in when all doctors are occupied—for example, in a crisis [66]. The use of smart technology in this sector is vital for saving lives during pandemics like COVID-19. The wearable monitoring devices capture and send data to a database for a doctor can analyse and then diagnose the patient or send a prescription.

Patients can be fitted with smart pills and smart bands (IoT) that monitor and collect specific data to feed a database during pandemics. These devices help doctors and other machines (machine learning) to learn disease patterns and symptoms, giving doctors a chance to understand symptoms and analyse the symptoms to develop quick and safe diagnostics [67]. During times of quarantine, such strategies can enhance safety for both the patient and health practitioners as machine learning technology prevents physical contact with patients infected with deadly airborne viruses.

Cloud computing is also an efficient part of the IoT sector. It helps to connect a wide variety of machine learning AI devices to understand data through analysis and storage. Another important feature of cloud computing is that it can store a huge amount of data and, therefore, sustain the needs of the healthcare system. Due to its data-sharing capabilities, cloud computing can also allow different devices to access the information. On the other hand, cloud computing currently faces some challenges that need to be addressed. These challenges could open up new research opportunities for scientists and researchers seeking to improve ML and IoT's usability in the healthcare industry. One of these challenges is data privacy and security. Medical records in the healthcare industry are highly sensitive and need to be carefully protected as they contain individuals' protected health information (PHI). Therefore, strict regulations, such as the Health Insurance Portability and Accountability Act (HIPAA) [72], have been introduced to regulate the process of accessing and analysing these data. This creates a significant challenge for modern data mining and ML technologies, such as deep learning, which typically require a large amount of training data. Sharing this type of sensitive information to improve quality-of-care delivery can compromise patient privacy. Several solutions for preserving patient privacy with ML technology have been introduced.

One solution is called federated learning (FL). This new ML paradigm uses deep learning to train and enable mobile devices and servers to build a common, robust ML model without sharing data [73]. FL also enables researchers to address critical issues such as data security, data access rights, and heterogeneous data access. Storing data in a centralised cloud computing is an additional issue for ML because using the same server to collect shared information from different devices and maintaining a generic model can make the server vulnerable to server malfunction and bias. This might also result in having an inaccurately trained model that will negatively influence the accuracy of the predicted outcome. Therefore, decentralised data storage is currently one of the best practices. One technology that has decentralised data storage capabilities is blockchain.

Table 4. Summary of the reviewed applications.

Reference	Application Name	Brief Description	ML Algorithms	Issues Addressed	Current Challenges	Future Work	Comparison with Existing Reviews
[45]	Medical imaging	Medical imaging is largely manual today as it entails a health professional examining images to determine abnormalities. However, machine learning algorithms can be used to automate this process and enhance the accuracy of the imaging process.	Artificial neural networks (ANNs) and convolutional neural networks (CNNs)	The use of machine learning addresses the issues of accuracy and efficiency when imaging is done manually.	High dependency on the quality and amount of training sets. Ethical and legal issues concerning the use of ML in healthcare. It is often difficult to explain the outputs of deep learning techniques logically.	Improving the quality of training datasets to improve accuracy and patient-centredness.	Existing reviews on medical imaging, such as [68], published in 2019, focus on providing a broad overview of the advances being made in this area of ML. The proposed review intends to offer an updated assessment of medical imaging ML algorithms and their application.
[47]	Diagnosis of diseases	Clinical diagnosis can benefit from machine learning by improving the quality and efficiency of decision-making.	Image-based deep learning	Wrong patient diagnoses result in inappropriate interventions and adverse outcomes.	The lack of sound laws and regulations defining the utilization of ML in healthcare. Obtaining well-annotated data for supervised learning is challenging.	Integrating ML into electronic medical records to support timely and accurate disease diagnoses.	Reviews on clinical diagnoses tend to focus on a specific disease type or group. For example, Schaefer et al. (2020) focused on rare diseases [52]. The proposed review hopes to continue in this vein but add on an in-depth examination of the application of these ML algorithms in practical environments and the potential benefits.

Table 4. Cont.

Reference	Application Name	Brief Description	ML Algorithms	Issues Addressed	Current Challenges	Future Work	Comparison with Existing Reviews
[49]	Behaviour modification or treatment	The integration of machine learning into behavioural change programs can help with determining what works and what does not.	Various machine learning and reasoning methods, including natural language processing	The inability to synthesize and deliver evidence on behavioural change interventions to user need and context to improve the usefulness of evidence.	The lack of a behavioural change intervention knowledge system consisting of an ontology, process, and resources for annotating reports, an automated annotator, ML and reasoning algorithms, and user interface.	The utilization of evidence from machine learning programs to guide behavioural change interventions.	Based on the researcher's exploration, there are no reviews systematically examining behavioural modification or treatment machine learning algorithms. Accordingly, these applications of ML ought to be assessed.
[51]	Clinical trial research	There is a need to develop machine learning algorithms capable for continual learning from clinical data.	Deep learning techniques	The difficulty of drawing insights from vast amounts of clinical data using human capabilities.	The problem of utilizing deep learning models on complex medical datasets. The need for high volumes of well-labelled training datasets. Ethical issues surrounding machine learning.	The continued collection of training datasets to improve the applicability of deep learning in clinical research trials.	Some review studies in this area exist. For example, Zame et al. (2020) reviewed the application of ML in clinical trials in the current COVID-19 setting [52]. The proposed research will examine ML applications in different clinical trial efforts.

Table 4. Cont.

Reference	Application Name	Brief Description	ML Algorithms	Issues Addressed	Current Challenges	Future Work	Comparison with Existing Reviews
[53]	Smart electronic health records	The inclusion of machine learning in electronic health records creates smart systems with the capability to perform disease diagnosis, progression prediction, and risk assessment.	Deep learning, natural language processing, and supervised machine learning	Current electronic health records store clinical data but do not support clinical decision-making.	Preparing data before they are fed into a machine learning algorithm remains a challenging task. Additionally, it is difficult to incorporate patient-specific factors in machine learning models.	The widespread adoption of smart electronic health records to support the management of different conditions or diseases.	Reviews assessing the incorporation of ML into electronic health records are few. Shinozaki (2020) reviewed the inclusion of ML in electronic health records to aid drug development [69]. However, additional research is required to determine whether ML can help further patient-centred care, improve the quality of care, and enhance efficiency. The proposed review hopes to address these components.
[56]	Epidemic outbreak prediction	Disease surveillance can benefit from machine learning as it allows for the prediction of epidemics, hence enabling the implementation of appropriate safeguards.	Deep neural network (DNN), long short-term memory (LSTM) learning, and the autoregressive integrated moving average (ARIMA)	The difficulty of preparing for and dealing with infectious diseases due to a lack of knowledge or forecasts.	The low accuracy of predictive models. The challenge of choosing parameters to utilize with the machine learning models.	The use of predictive models to forecast a range of infectious diseases.	There have been studies reviewing the application of ML to disease outbreak prediction. Philemon et al. (2019) reviewed the utilization of ANN to predict outbreaks [70]. The proposed review intends to further the review efforts and examine ways of enhancing the predictive accuracy of the ML algorithms created.

Table 4. Cont.

Reference	Application Name	Brief Description	ML Algorithms	Issues Addressed	Current Challenges	Future Work	Comparison with Existing Reviews
[58]	Heart disease prediction	AI can be utilized to predict heart disease, hence enabling patients and health providers to implement preventive measures.	Deep learning and artificial neural networks	There is a need for accurate prediction of cardiovascular diseases, as well as the implementation of effective treatments to improve patient outcomes.	The lack of ethical guidelines to direct the adoption of heart disease prediction algorithms. Machine learning algorithms cannot solve highly abstract reasoning problems.	Extending the utilization of ML in clinical decision-making to include patient-centred predictive analytics.	In this area, current and comprehensive reviews have been performed [71]. Therefore, the objective of the planned review is to report the findings and suggest areas of further research.
[60]	Diagnostic and prognostic models for COVID-19	This study examined the prediction models for COVID-19 and found that they are poorly designed.	Deep learning models	The need to review prediction models for the diagnosis and prognosis of COVID-19 to support their use to guide decision-making.	The developed models are problematic due to the poor training datasets used.	Collect high-volume and quality datasets to train COVID-19 prediction models.	Systematic reviews assessing the application of ML for COVID-19 diagnosis and prognosis have been conducted [60]. As knowledge of the disease continues to improve, better and more effective models can emerge, hence the need for continual reviews.
[63]	Personalized care	Machine learning algorithms provide an avenue for offering person-centred care.	Deep neural networks, deep learning, supervised and unsupervised learning, and many others	The inability to provide personalized care despite the increasing accumulation of personal data.	There is a need for the continued accumulation of high-quality training datasets.	Creating systems that can be integrated into electronic health records to promote personalized medicine.	Fröhlich et al. (2018) reviewed the application of ML to enable personalized care. The study identified some of the challenges associated with this endeavor. As such, the proposed review hopes to explore how subsequent studies have addressed these challenges.

There are devices capable of monitoring body temperature, blood pressure, and heart rate. They are useful for collecting and storing data about patients and hence can contribute to diagnosis. IoT and machine learning can help keep healthcare professionals abreast of changes, which is important for a healthy society. The storage of diagnostic data and COVID-19 symptoms is key to ensuring that a disease is wiped out or a vaccine is found since data can be stored in a central database and accessed by scientists and medical practitioners for cross-examination, analysis, and real-time sharing of results.

6. Conclusions

The healthcare sector is one of the most complex in terms of the level of responsibility and strict regulations, which makes it an important and vital sector for innovations. The Internet of things (IoT) has opened up a world of possibilities in the healthcare sector and could be the solution to many problems. Applying the medical IoT will bring about great opportunities for telemedicine, remote monitoring of patients' condition, and much more. This could be possible with the help of ML models. In this article, we summarised the most powerful ML algorithms, listed some ML applications in the healthcare field, and analysed IoT and machine learning in the healthcare system to predict future trends.

Author Contributions: Conceptualization, A.A. and B.A.; methodology, A.A.; formal analysis, A.A., B.A. and W.H.; investigation, A.A., B.A.; resources, B.A.; writing—original draft preparation, A.A., B.A.; writing—review and editing, A.A., B.A., W.H.; supervision, W.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors would like to acknowledge that the journal Forecasting waives the APC of the article, thanks to the journal.

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

References

1. Mtonga, K.; Kumaran, S.; Mikeka, C.; Jayavel, K.; Nsenga, J. Machine Learning-Based Patient Load Prediction and IoT Integrated Intelligent Patient Transfer Systems. *Future Internet* **2019**, *11*, 236. [[CrossRef](#)]
2. Mosenia, A.; Sur-Kolay, S.; Raghunathan, A.; Jha, N.K. Wearable Medical Sensor-Based System Design: A Survey. *IEEE Trans. Multi-Scale Comput. Syst.* **2017**, *3*, 124–138. [[CrossRef](#)]
3. Iqbal, N.; Jamil, F.; Ahmad, S.; Kim, D. A Novel Blockchain-Based Integrity and Reliable Veterinary Clinic Information Management System Using Predictive Analytics for Provisioning of Quality Health Services. *IEEE Access* **2021**, *9*, 8069–8098. [[CrossRef](#)]
4. Wu, T.; Wu, F.; Redoute, J.-M.; Yuce, M.R. An Autonomous Wireless Body Area Network Implementation towards IoT Connected Healthcare Applications. *IEEE Access* **2017**, *5*, 11413–11422. [[CrossRef](#)]
5. Birje, M.N.; Hanji, S.S. Internet of things based distributed healthcare systems: A review. *J. Data Inf. Manag.* **2020**, *2*, 149–165. [[CrossRef](#)]
6. Shahbazi, Z.; Byun, Y.-C. Towards a Secure Thermal-Energy Aware Routing Protocol in Wireless Body Area Network Based on Blockchain Technology. *Sensors* **2020**, *20*, 3604. [[CrossRef](#)]
7. Kumar, P.M.; Gandhi, U.D. A novel three-tier Internet of Things architecture with machine learning algorithm for early detection of heart diseases. *Comput. Electr. Eng.* **2018**, *65*, 222–235. [[CrossRef](#)]
8. Yuvaraj, N.; SriPreethaa, K.R. Diabetes prediction in healthcare systems using machine learning algorithms on Hadoop cluster. *Clust. Comput.* **2017**, *22*, 1–9. [[CrossRef](#)]
9. Miotto, R.; Wang, F.; Wang, S.; Jiang, X.; Dudley, J.T. Deep learning for healthcare: Review, opportunities and challenges. *Brief. Bioinform.* **2018**, *19*, 1236–1246. [[CrossRef](#)]
10. Shahbazi, Z.; Byun, Y.-C. Improving Transactional Data System Based on an Edge Computing–Blockchain–Machine Learning Integrated Framework. *Processes* **2021**, *9*, 92. [[CrossRef](#)]
11. Sohaib, O.; Lu, H.; Hussain, W. Internet of Things (IoT) in E-commerce: For people with disabilities. In Proceedings of the 2017 12th IEEE Conference on Industrial Electronics and Applications (ICIEA), Siem Reap, Cambodia, 18–20 June 2017; pp. 419–423.
12. Gao, H.; Qin, X.; Barroso, R.J.D.; Hussain, W.; Xu, Y.; Yin, Y. Collaborative Learning-Based Industrial IoT API Recommendation for Software-Defined Devices: The Implicit Knowledge Discovery Perspective. *IEEE Trans. Emerg. Top. Comput. Intell.* **2020**, 1–11. [[CrossRef](#)]
13. Rajkomar, A.; Hardt, M.; Howell, M.D.; Corrado, G.; Chin, M.H. Ensuring Fairness in Machine Learning to Advance Health Equity. *Ann. Intern. Med.* **2018**, *169*, 866–872. [[CrossRef](#)]

14. Jamil, F.; Ahmad, S.; Iqbal, N.; Kim, D.-H. Towards a remote monitoring of patient vital signs based on IoT-based blockchain integrity management platforms in smart hospitals. *Sensors* **2020**, *20*, 2195. [[CrossRef](#)] [[PubMed](#)]
15. Jadhav, S.; Kasar, R.; Lade, N.; Patil, M.; Kolte, S. Disease Prediction by Machine Learning from Healthcare Communities. *Int. J. Sci. Res. Sci. Technol.* **2019**, 29–35. [[CrossRef](#)]
16. Ahmad, M.A.; Eckert, C.; Teredesai, A. Interpretable Machine Learning in Healthcare. In Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, Washington, DC, USA, 24–29 August 2018; pp. 559–560.
17. Panch, T.; Szolovits, P.; Atun, R. Artificial intelligence, machine learning and health systems. *J. Glob. Health* **2018**, *8*, 020303. [[CrossRef](#)] [[PubMed](#)]
18. Jamil, F.; Hang, L.; Kim, K.; Kim, D. A Novel Medical Blockchain Model for Drug Supply Chain Integrity Management in a Smart Hospital. *Electronics* **2019**, *8*, 505. [[CrossRef](#)]
19. Wiens, J.; Shenoy, E.S. Machine Learning for Healthcare: On the Verge of a Major Shift in Healthcare Epidemiology. *Clin. Infect. Dis.* **2018**, *66*, 149–153. [[CrossRef](#)] [[PubMed](#)]
20. Hung, C.-Y.; Chen, W.-C.; Lai, P.-T.; Lin, C.-H.; Lee, C.-C. Comparing deep neural network and other machine learning algorithms for stroke prediction in a large-scale population-based electronic medical claims database. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Jeju Island, Korea, 11–15 July 2017; pp. 3110–3113.
21. Ngiam, K.Y.; Khor, W. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol.* **2019**, *20*, e262–e273. [[CrossRef](#)]
22. Dike, H.U.; Zhou, Y.; Deveerasetty, K.K.; Wu, Q. Unsupervised Learning Based on Artificial Neural Network: A Review. In Proceedings of the 2018 IEEE International Conference on Cyborg and Bionic Systems (CBS), Shenzhen, China, 25–27 October 2018; pp. 322–327.
23. Osisanwo, F.; Akinsola, J.; Awodele, O.; Hinmikaiye, J.; Olakanmi, O.; Akinjobi, J. Supervised machine learning algorithms: Classification and comparison. *Int. J. Comput. Trends Technol.* **2017**, *48*, 128–138.
24. Praveena, M.; Jaiganesh, V. A Literature Review on Supervised Machine Learning Algorithms and Boosting Process. *Int. J. Comput. Appl.* **2017**, *169*, 32–35. [[CrossRef](#)]
25. Friedman, J.; Hastie, T.; Tibshirani, R. *The Elements of Statistical Learning (No. 10)*; Springer Series in Statistics: New York, NY, USA, 2001.
26. Golub, G.H.; Heath, M.; Wahba, G. Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter. *Technometrics* **1979**, *21*, 215. [[CrossRef](#)]
27. Tambuskar, D.; Narkhede, B.; Mahapatra, S.S. A flexible clustering approach for virtual cell formation considering real-life production factors using Kohonen self-organising map. *Int. J. Ind. Syst. Eng.* **2018**, *28*, 193–215. [[CrossRef](#)]
28. Ben Ali, J.; Saidi, L.; Harrath, S.; Bechhoefer, E.; Benbouzid, M. Online automatic diagnosis of wind turbine bearings progressive degradations under real experimental conditions based on unsupervised machine learning. *Appl. Acoust.* **2018**, *132*, 167–181. [[CrossRef](#)]
29. Van Engelen, J.E.; Hoos, H.H. A survey on semi-supervised learning. *Mach. Learn.* **2019**, *109*, 373–440. [[CrossRef](#)]
30. Hussain, W.; Sohaib, O. Analysing Cloud QoS Prediction Approaches and Its Control Parameters: Considering Overall Accuracy and Freshness of a Dataset. *IEEE Access* **2019**, *7*, 82649–82671. [[CrossRef](#)]
31. Hussain, W.; Hussain, F.K.; Hussain, O.K.; Chang, E. Provider-Based Optimized Personalized Viable SLA (OPV-SLA) Framework to Prevent SLA Violation. *Comput. J.* **2016**, *59*, 1760–1783. [[CrossRef](#)]
32. Hussain, W.; Sohaib, O.; Naderpour, M.; Gao, H. Cloud Marginal Resource Allocation: A Decision Support Model. *Mob. Netw. Appl.* **2020**, *25*, 1418–1433. [[CrossRef](#)]
33. Hussain, W.; Hussain, F.K.; Hussain, O.; Bagia, R.; Chang, E. Risk-based framework for SLA violation abatement from the cloud service provider’s perspective. *Comput. J.* **2018**, *61*, 1306–1322. [[CrossRef](#)]
34. Kaur, G.; Oberoi, A. Novel Approach for Brain Tumor Detection Based on Naïve Bayes Classification. In *Data Management, Analytics and Innovation*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 451–462.
35. Singh, A.; Thakur, N.; Sharma, A. A review of supervised machine learning algorithms. In Proceedings of the 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 16–18 March 2016; pp. 1310–1315.
36. Oshiro, T.M.; Perez, P.S.; Baranauskas, J.A. How Many Trees in a Random Forest? In *International Workshop on Machine Learning and Data Mining in Pattern Recognition*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 154–168.
37. Ye, J.; Chow, J.-H.; Chen, J.; Zheng, Z. Stochastic gradient boosted distributed decision trees. In Proceedings of the 18th ACM conference on Information and knowledge management—CIKM’09, Hong Kong, China, 2–6 November 2009; pp. 2061–2064.
38. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
39. Noble, W.S. What is a support vector machine? *Nat. Biotechnol.* **2006**, *24*, 1565–1567. [[CrossRef](#)]
40. Bisong, E. *Building Machine Learning and Deep Learning Models on Google Cloud Platform*; Springer: Berlin/Heidelberg, Germany, 2019.
41. Olmedo, M.T.C.; Paegelow, M.; Mas, J.-F.; Escobar, F. *Geomatic Approaches for Modeling Land Change Scenarios*; Springer: Berlin/Heidelberg, Germany, 2018.

42. Chauhan, R.; Ghanshala, K.K.; Joshi, R.C. Convolutional Neural Network (CNN) for Image Detection and Recognition. In Proceedings of the 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), Kolkata, India, 22–23 November 2018; pp. 278–282.
43. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6.
44. Aliberti, A.; Bagatin, A.; Acquaviva, A.; Macii, E.; Patti, E. Data Driven Patient-Specialized Neural Networks for Blood Glucose Prediction. In Proceedings of the 2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), London, UK, 6–10 July 2020; pp. 1–6.
45. Kim, M.; Yun, J.; Cho, Y.; Shin, K.; Jang, R.; Bae, H.-J.; Kim, N. Deep Learning in Medical Imaging. *Neurospine* **2019**, *16*, 657–668. [[CrossRef](#)]
46. Desai, S.B.; Pareek, A.; Lungren, M.P. Deep learning and its role in COVID-19 medical imaging. *Intell. Med.* **2020**, *3*, 100013. [[CrossRef](#)] [[PubMed](#)]
47. Xu, J.; Xue, K.; Zhang, K. Current status and future trends of clinical diagnoses via image-based deep learning. *Theranostics* **2019**, *9*, 7556–7565. [[CrossRef](#)]
48. Battineni, G.; Sagaro, G.G.; Chinatalapudi, N.; Amenta, F. Applications of Machine Learning Predictive Models in the Chronic Disease Diagnosis. *J. Pers. Med.* **2020**, *10*, 21. [[CrossRef](#)] [[PubMed](#)]
49. Michie, S.; Thomas, J.; John, S.-T.; Mac Aonghusa, P.; Shawe-Taylor, J.; Kelly, M.P.; Deleris, L.A.; Finnerty, A.N.; Marques, M.M.; Norris, E.; et al. The Human Behaviour-Change Project: Harnessing the power of artificial intelligence and machine learning for evidence synthesis and interpretation. *Implement. Sci.* **2017**, *12*, 1–12. [[CrossRef](#)] [[PubMed](#)]
50. Cahyadi, A.; Razak, A.; Abdillah, H.; Junaedi, F.; Taligansing, S.Y. Machine Learning Based Behavioral Modification. *Int. J. Eng. Adv. Technol.* **2019**, *8*, 1134–1138.
51. Shah, P.; Kendall, F.; Khozin, S.; Goosen, R.; Hu, J.; Laramie, J.; Ringel, M.; Schork, N. Artificial intelligence and machine learning in clinical development: A translational perspective. *NPJ Digit. Med.* **2019**, *2*, 1–5. [[CrossRef](#)]
52. Zame, W.R.; Bica, I.; Shen, C.; Curth, A.; Lee, H.-S.; Bailey, S.; Weatherall, J.; Wright, D.; Bretz, F.; Van Der Schaar, M. Machine learning for clinical trials in the era of COVID-19. *Stat. Biopharm. Res.* **2020**, *12*, 506–517. [[CrossRef](#)]
53. Lin, W.-C.; Chen, J.S.; Chiang, M.F.; Hribar, M.R. Applications of Artificial Intelligence to Electronic Health Record Data in Ophthalmology. *Transl. Vis. Sci. Technol.* **2020**, *9*, 13. [[CrossRef](#)] [[PubMed](#)]
54. Wang, S.; Pathak, J.; Zhang, Y. Using Electronic Health Records and Machine Learning to Predict Postpartum Depression. *Stud. Health Technol. Inform.* **2019**, *264*, 888–892.
55. Solares, J.R.A.; Raimondi, F.E.D.; Zhu, Y.; Rahimian, F.; Canoy, D.; Tran, J.; Gomes, A.C.P.; Payberah, A.H.; Zottoli, M.; Nazarzadeh, M.; et al. Deep learning for electronic health records: A comparative review of multiple deep neural architectures. *J. Biomed. Inform.* **2020**, *101*, 103337. [[CrossRef](#)]
56. Chae, S.; Kwon, S.; Lee, D. Predicting Infectious Disease Using Deep Learning and Big Data. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1596. [[CrossRef](#)]
57. Ibrahim, N.; Akhir, N.S.M.; Hassan, F.H. Predictive analysis effectiveness in determining the epidemic disease infected area. *AIP Conf. Proc.* **2017**, *1891*, 20064. [[CrossRef](#)]
58. Yan, Y.; Zhang, J.-W.; Zang, G.-Y.; Pu, J. The primary use of artificial intelligence in cardiovascular diseases: What kind of potential role does artificial intelligence play in future medicine? *J. Geriatr. Cardiol. JGC* **2019**, *16*, 585–591.
59. Alm Mustafa, K.M. Prediction of heart disease and classifiers' sensitivity analysis. *BMC Bioinform.* **2020**, *21*, 278. [[CrossRef](#)] [[PubMed](#)]
60. Wynants, L.; Van Calster, B.; Collins, G.S.; Riley, R.D.; Heinze, G.; Schuit, E.; Bonten, M.M.J.; Dahly, D.L.; Damen, J.A.; Debray, T.P.A.; et al. Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. *BMJ* **2020**, *369*, m1328. [[CrossRef](#)] [[PubMed](#)]
61. Zoabi, Y.; Deri-Rozov, S.; Shomron, N. Machine learning-based prediction of COVID-19 diagnosis based on symptoms. *NPJ Digit. Med.* **2021**, *4*, 1–5. [[CrossRef](#)]
62. Wilkinson, J.; Arnold, K.F.; Murray, E.J.; van Smeden, M.; Carr, K.; Sippy, R.; de Kamps, M.; Beam, A.; Konigorski, S.; Lippert, C.; et al. Time to reality check the promises of machine learning-powered precision medicine. *Lancet Digit. Health* **2020**, *2*, e677–e680. [[CrossRef](#)]
63. Ahmed, Z.; Mohamed, K.; Zeeshan, S.; Dong, X. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. *Database J. Biol. Databases Curation* **2020**, *2020*. [[CrossRef](#)] [[PubMed](#)]
64. Vanani, I.R.; Amirhosseini, M. IoT-Based Diseases Prediction and Diagnosis System for Healthcare. In *Internet of Things for Healthcare Technologies*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 21–48.
65. Wang, M.-H.; Chen, H.-K.; Hsu, M.-H.; Wang, H.-C.; Yeh, Y.-T. Cloud Computing for Infectious Disease Surveillance and Control: Development and Evaluation of a Hospital Automated Laboratory Reporting System. *J. Med. Internet Res.* **2018**, *20*, e10886. [[CrossRef](#)] [[PubMed](#)]
66. Shahzad, A.; Lee, Y.S.; Lee, M.; Kim, Y.-G.; Xiong, N. Real-Time Cloud-Based Health Tracking and Monitoring System in Designed Boundary for Cardiology Patients. *J. Sens.* **2018**, *2018*, 1–15. [[CrossRef](#)]
67. Mocnik, F.-B.; Raposo, P.; Feringa, W.; Kraak, M.-J.; Köbben, B. Epidemics and pandemics in maps—the case of COVID-19. *J. Maps* **2020**, *16*, 144–152. [[CrossRef](#)]

68. Lundervold, A.S.; Lundervold, A. An overview of deep learning in medical imaging focusing on MRI. *Z. Med. Phys.* **2019**, *29*, 102–127. [[CrossRef](#)] [[PubMed](#)]
69. Shinozaki, A. Electronic Medical Records and Machine Learning in Approaches to Drug Development. In *Artificial Intelligence in Oncology Drug Discovery and Development*; IntechOpen: London, UK, 2020; p. 51.
70. Datilo, P.M.; Ismail, Z.; Dare, J. A Review of Epidemic Forecasting Using Artificial Neural Networks. *Int. J. Epidemiol. Res.* **2019**, *6*, 132–143. [[CrossRef](#)]
71. Yahaya, L.; Oye, N.D.; Garba, E.J. A Comprehensive Review on Heart Disease Prediction Using Data Mining and Machine Learning Techniques. *Am. J. Artif. Intell.* **2020**, *4*, 20. [[CrossRef](#)]
72. Alharbi, R.; Almagwashi, H. The Privacy requirements for wearable IoT devices in healthcare domain. In Proceedings of the 2019 7th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW), Istanbul, Turkey, 26–28 August 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 18–25.
73. Xu, J.; Glicksberg, B.S.; Su, C.; Walker, P.; Bian, J.; Wang, F. Federated Learning for Healthcare Informatics. *J. Health Inform. Res.* **2021**, *5*, 1–19. [[CrossRef](#)] [[PubMed](#)]