

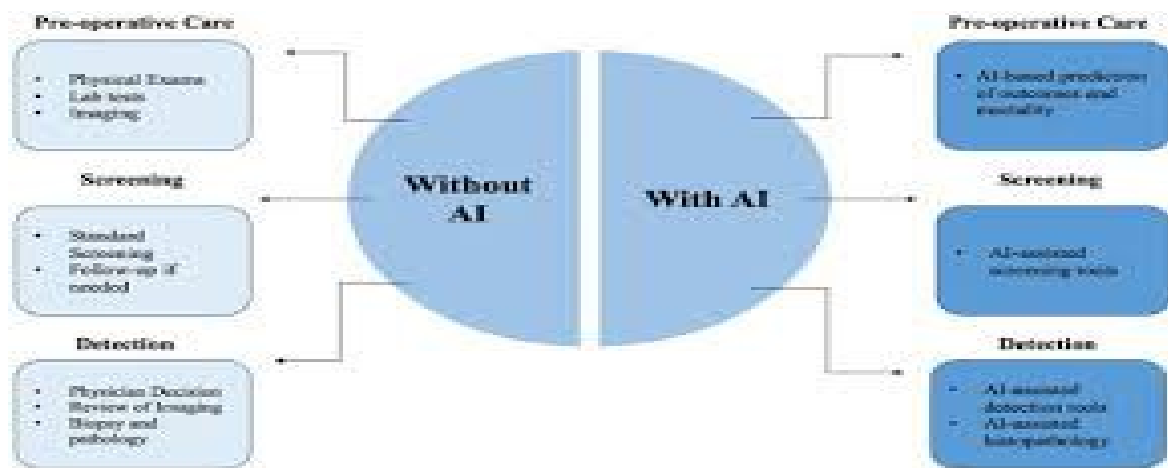
Healthcare outcom **prediction project**

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



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 The IoT presents possible measures to decreases 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 SriPreethaa 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-electromechanical 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 Forecasting 2021, 3 183 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

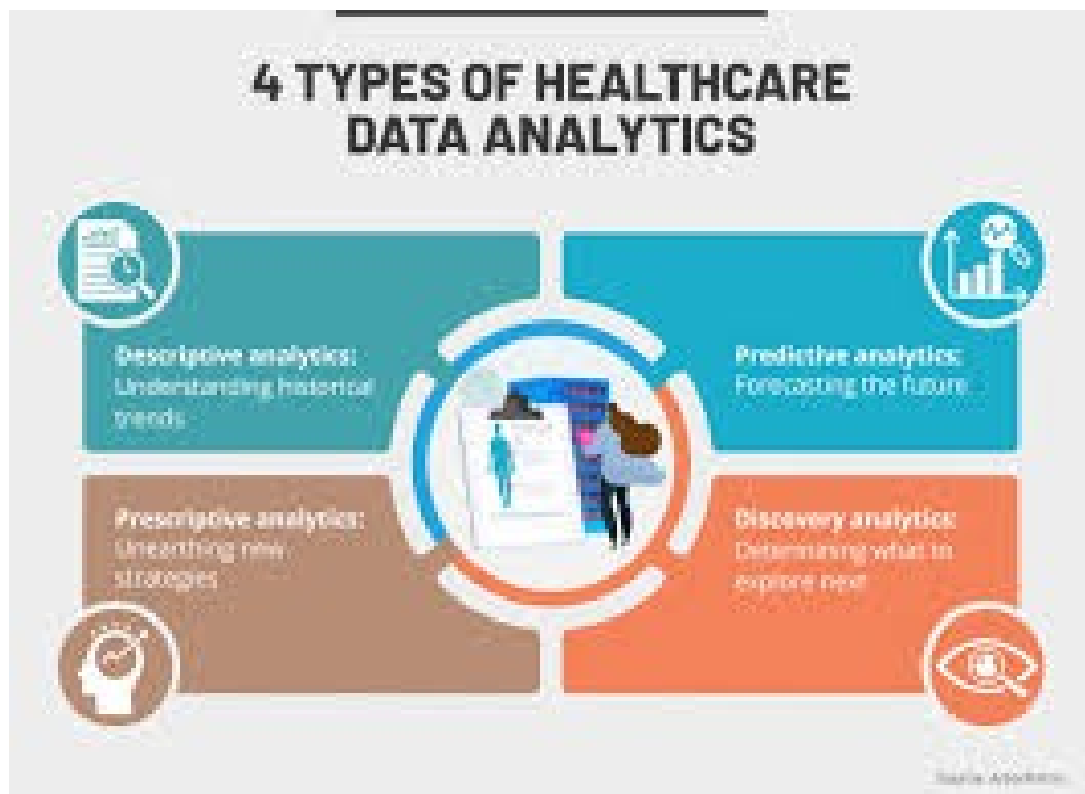
store the information and enhance data analysis [17]. Various open-source cloud computing platforms are compatible with the Raspbian Jessie and Raspberry Pi board [18]. These devices utilize machine learning algorithms to assess the stored data to recognize the existence of any anomalies [19]. Therefore, 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. 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. Forecasting 2021, 3 185
Forecasting 2021, 3 FOR PEER REVIEW 5 Figure 1. Machine learning algorithms' classification.

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

Figure 1. Machine learning algorithms' classification.

Machine learning algorithms' classification. 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. Figure 2.

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

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 Forecasting 2021, 3 186 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.

Forecasting 2021, 3 187 Forecasting 2021, 3 FOR PEER REVIEW 7 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 crossvalidation, 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

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

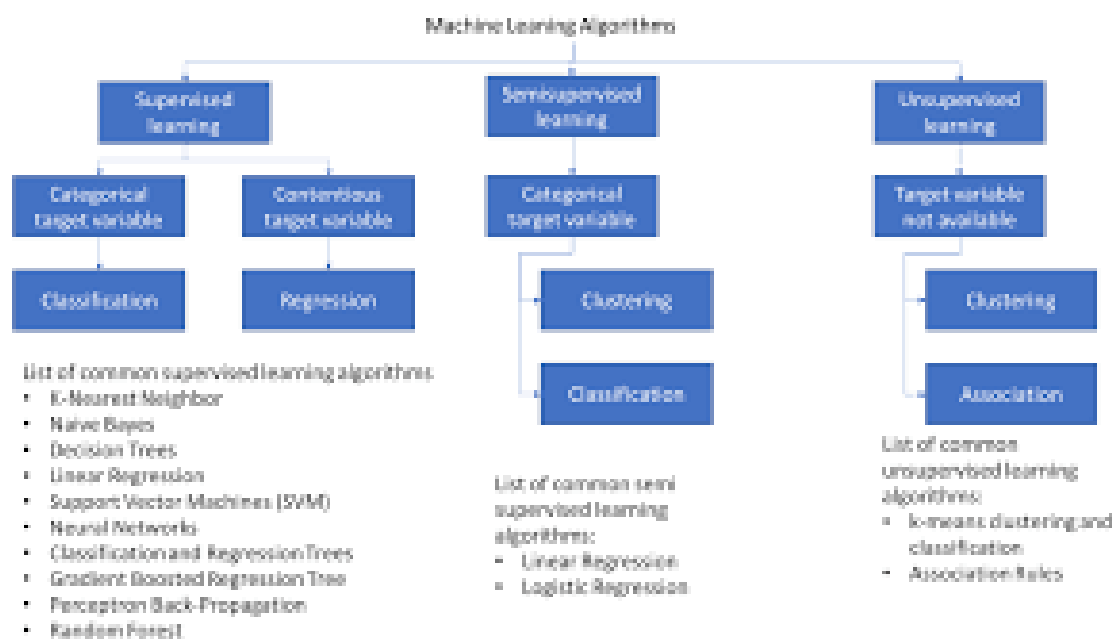
procedure is timeconsuming, which might hinder creating completely labelled training and bring about heavy expenses. As a result, in certain instances, semisupervised learning may prove to be a better solution. In the case of a limited number of labelled samples, semisupervised learning is mostly used 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].

Forecasting 2021, 3 188

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

Learning Type	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 SemiSupervised
 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

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 2021, 3 190 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.

$$P(A|B) = P(B|A) * P(A) / P(B)$$

3.3 Decision Tree (DT) Classification The simple classification algorithm consisting of the internal node and one classlabelled 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

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 Forecasting 2021, 3 192 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.

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

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.

4.2Diagnosis 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

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.

4.3. 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.

4.4. 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 Forecasting 2021, 3 197 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.

4.5. 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.

4.6. 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.

4.7. 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 decisionmaking by allowing healthcare professionals to provide personalized care. The input data Forecasting 2021, 3 198

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

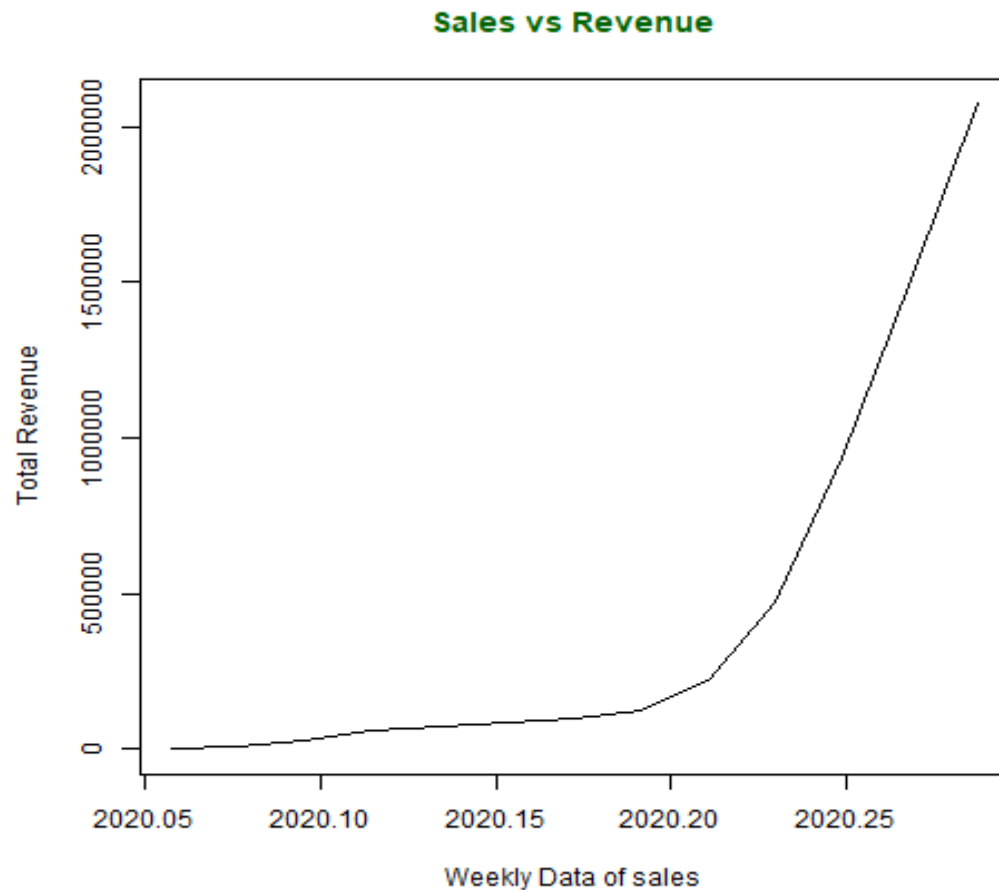
machine learning AI devices to understand data through analysis and storage. Another important feature of cloud computing is that it can store a huge amount of data and, therefore, sustain the needs of the healthcare system. Due to its data-sharing capabilities, cloud computing can also allow different 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

Coding:

```
x <- c(580, 7813, 28266, 59287, 75700,  
87820, 95314, 126214, 218843, 471497,  
936851, 1508725, 2072113)  
  
# library required for decimal_date() function  
library(lubridate)  
  
# output to be created as png file  
png(file = "predictiveAnalysis.png")  
  
# creating time series object  
# from date 22 January, 2020  
mts <- ts(x, start = decimal_date(ymd("2020-01-22")),  
frequency = 365.25 / 7)  
  
# plotting the graph  
plot(mts, xlab = "Weekly Data of sales",  
ylab = "Total Revenue",  
main = "Sales vs Revenue",  
col.main = "darkgreen")  
  
# saving the file  
dev.off()
```

Output:



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