IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, October 2026
Impact Factor: 8.2

Healthcare Predictive Analysis Using Machine Learning Techniques and Classification Methods.

Ms. Shiwani P. Mendhe (Bhure) and Dr. Aruna J. Chamatkar
Research Scholar, Department of Electronics & Computer Science, R.T.M.N.U. Nagpur, Maharashtra, India
Associate Professor and Head, Department of Computer Science, Kamla Nehru College, Nagpur, Maharashtra, India
shivanimendhe18@gmail.com and aruna.ayush1007@gmail.com

Abstract: *This paper explores the increasing relevance of machine learning with respect to healthcare system research. It was extremely important to predict the outcomes of healthcare in recent years to save lives. Intelligent systems, which are able to analyze the connections between complex data and use these connections for further prediction, are being actively developed in the sphere of healthcare. Thus, artificial intelligence makes significant changes in healthcare, and machine learning and deep learning systems become very important in the development of diagnostic and predicting techniques. Predictive analytics becomes an extremely important necessity in healthcare, because it is able to improve the accuracy of predictions, which can save lives when they are accurate and made in time. The primary aim of this research is to find the best successful algorithms for using machine learning in healthcare systems and predictive analytics. In addition, there is the introduction of machine learning classification methods and techniques.*

Keywords: Machine Learning (ML), Deep Learning (DL), Classification Methods, Healthcare Forecasting, Predictive Techniques.

I. INTRODUCTION

Increasing applications of machine learning in the healthcare sector provides an indication of a future wherein data, analysis, and innovation work together to help many people, even though they are unaware of the process. It is expected that such applications driven by machine learning will soon be combined real time data collected from various healthcare organizations in several countries.

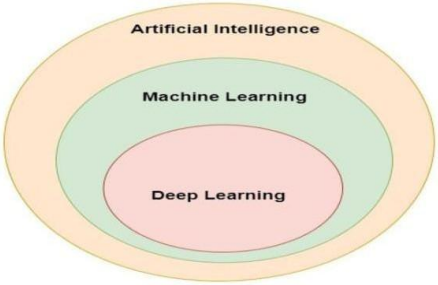
Health care includes a wide array of processes that deal with the diagnosis, treatment, and prevention of diseases. The medical industry in many nations is developing very rapidly. Data is abundant in the health care sector since it generates large quantities of data, such as health records.

The field of health informatics is transforming into an extremely research-based field and becoming the primary beneficiary of funding from the public sector. With the advent of computers and algorithms, there have been many technological advancements in the field of healthcare, which makes it important for the industry to embrace all these technologies. Thus, the merging of healthcare with computing has created health informatics. This merging of fields is expected to make healthcare more efficient and effective, along with improving the quality of healthcare delivery and decreasing costs. Machine learning algorithms will improve the prediction of various healthcare problems significantly. Health care is the joint effort of society to promote and provide health care.

There has been seen a significant change in the twentieth century concerning the concepts for health and the prevention of illness and disabilities. Healthcare delivery refers to the organized efforts either private or public to help people recover from sickness or disabilities. It can also be seen as a set of standardized norms which help in assessing actions and situations to make decisions. It is a complicated, multi-faceted process. Healthcare's primary objective is to diagnose and cure diseases and disabilities.

Copyright to IJARSCT
www.ijarsct.co.in
DOI: 10.48175/IJARSCT-40039

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, October 2026
Impact Factor: 8.2



Predictive analytics in healthcare plays an important part in health industry. It is the effectiveness of the process which greatly impacts how precise disease predictions will be made, possibly saving people's lives in the process of accurate and timely predictions.

However, wrong predictions might endanger the safety of patients. Thus, the urgent need for effective and reliable methods for healthcare predictive analysis arises. The objective of this research is to conduct a comprehensive study of the widely used methods of machine learning and deep learning in healthcare predictions.

Background

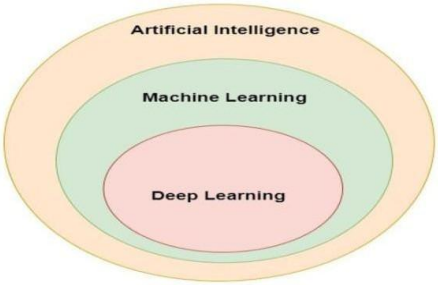
The vast research and development of advanced technology that Predictive analytics and deep learning techniques to predict health outcomes of individuals have gained greater interest in healthcare. Clinical prediction models have been instrumental in helping health care providers identify and manage patients who are likely to develop health disorders. Prediction models that take into account individual characteristics are used to predict health outcomes of patients and health care practices. Artificial intelligence (AI) is a system's ability to analyze data, using computers and technology to supplement human decision-making, issue resolution, and technical innovation.

The figure below shows that machine learning and deep learning are both branches of artificial intelligence.

Machine Learning

Technological advances in machine learning provide new approaches to developing learning models from complex data sets. This research article presents review of existing literature concerning application of machine learning technologies to enhance the healthcare sector. In contrast to various conventional prediction algorithms, machine learning algorithms achieve the highest level of prediction accuracy and exhibit convergence speed that surpasses all other methods for prediction. ML is based on ideas drawn from different disciplines such as computer science, statistics, and optimization. Essentially, all ML tasks can be seen as optimization problems in relation to a dataset. [2]

ML is based on ideas from numerous disciplines such as computer science, statistics, and optimization. Essentially, most ML questions can be reduced to an optimization problem based on the dataset. In recent years, machine learning has made significant contributions across multiple fields, particularly in areas such as computer vision. [11]

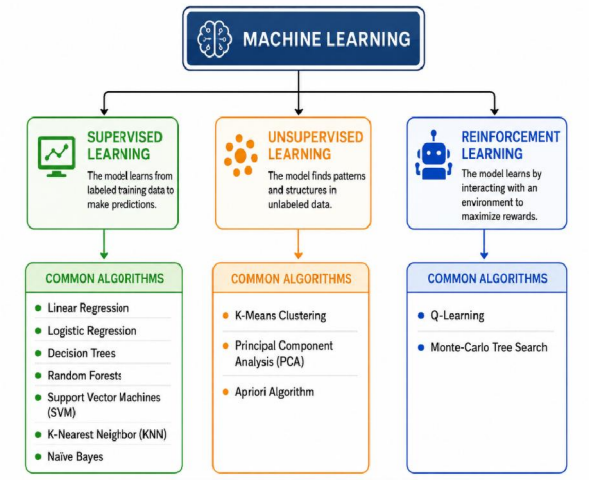


In above fig. traditional approach vs. machine learning (ML) approach. The starting point for the traditional approach to the analysis of data is the model being fed to the machine as input. For the ML (data-driven) approach, starting point is data, with the output being a model. [2]

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, October 2026
Impact Factor: 8.2

Machine Learning Classification Methods

Machine Learning (ML) employs different algorithms, methodologies, and processes to identify correlations within datasets and to develop both technologies for describing and prediction of these correlations. ML is generally connected with data mining, pattern recognition and deep learning.[14] Machine Learning algorithms can be divided into three categories: supervised learning, unsupervised learning, and reinforcement learning.



Supervised Learning

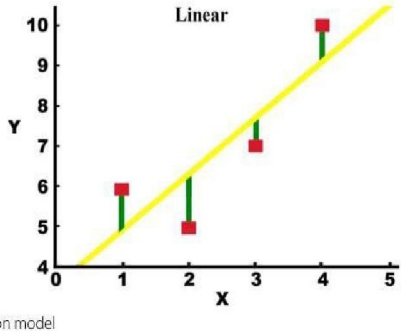
Supervised learning falls under the category of machine learning that attempts to analyze the association between input and output in a system, by taking into account a certain set of training examples that relate the inputs to their output. Methods which construct or create a mapping of the relationship between input and output variables through labeled data sets are referred to as supervised learning methods. [1] The model is trained through a labeled dataset, which is akin to how students learn basic mathematical operations from teachers.

Supervised learning requires the presence of labeled data, which contains the expected correct responses generated from the output of the algorithm. Examples of supervised learning include Linear Regression, Logistic Regression, Decision Trees, Random Forests, Support Vector Machines, K Nearest Neighbors, and Naïve Bayes.

Linear Regression

Linear Regression is a statistical method extensively used in predictive modeling. This is because it can predict the value of the dependent variable (Y) using the independent variable (X). The relationship between these two variables is represented by Equation, and both parameters should be continuous and numerical.

$$Y = m X + c$$



Linear regression model

Copyright to IJARSCT
www.ijarsct.co.in
DOI: 10.48175/IJARSCT-40039

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, October 2026
Impact Factor: 8.2

In the above equation, m indicates the slope, while c indicates the intercept. The equation helps infer the relationship between the independent variable (X) and dependent variable (Y). [12] As shown in Fig., data points have been highlighted in red, whereas random fluctuation (shown in green) indicate deviations from basic relation (showing in yellow) that exist between X and Y.

The core strength of linear regression lies in its simplicity, enabling the easy understanding and proper handling of overfitting by means of regularization techniques.[13] The core weakness of linear regression is its limitation to model nonlinear relationships. This is why the use of linear regression in most practical applications is generally not recommended due to oversimplification of complex situations. Popular programming languages used for implementing linear regression include Python, R, MATLAB, and Excel.

Logistic Regression

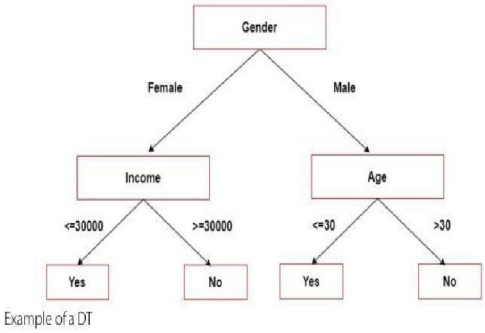
Logistic regression is often called a logistic model. It is a method that helps to analyze the connection between one or several independent variables and the categorical dependent variable by using probability of an event occurrence to fit the data into the logistic curve. The dependent variable should be a dichotomous variable, meaning that it can take two values only: true or false, 0 or 1, yes or no, or any other form of dichotomy. It is mainly used for prediction of the binary or multi-class response variables. [3]

The aim of using logistic regression analysis is prediction of categorical variables and dealing with categorization problems. [12] This regression technique can be used with different programming languages such as R, Python, Java, MATLAB, and others. The advantages of logistic regression include that it demonstrates the linear relation between dependent and independent variables, gives the best results and is easy to comprehend.[15] However, it has some drawbacks; it can predict numerical outputs only and is not good for non-linear data. Regression equation given as, [8]

$$p = \frac{1}{1+e^{-(b_0+b_1x_1+\cdots+b_nx_n)}}$$

Decision Tree

Decision Tree (DT) is a supervised learning model applied in classification tasks. Attribute values are organized according to their order, whether ascending or descending. [10] Being a tree-structured model, DT describes each path from the root by separating data in order to reach a Boolean conclusion on the leaf node. DT is a hierarchical model representing knowledge interactions using nodes and connections. Decision Tree Learning is one of the most popular classification techniques in use today. [10] In this technique, decision trees learned from the training data can be represented as a collection of if-then statements to enhance readability. [9] In classification models, nodes indicate specific targets. One example of a decision tree can be seen in Fig. Nonetheless, DT has several drawbacks such as being complicated due to expanding terminology, small changes leading to different structures, and time needed for data training. Popular tools for decision tree implementation include Python (Scikit-Learn), RStudio and Weka.



Example of a DT

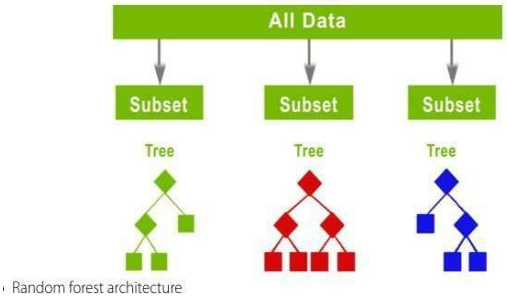
IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, October 2026
Impact Factor: 8.2

Random Forest

The Random Forest (RF) is another important technology that stands out with its reliable accuracy in many practical tasks. This approach can be used for classification as well as regression purposes. The method uses a number of Decision Trees (DTs) and combines their predictions. [10] Speaking about the RF classifier, with the increase of the number of DTs, accuracy usually increases. Consequently, RF constructs a set of DTs, which is called a forest, and unites them in order to enhance predictive power. The DTs are constructed from a random sample of the data and are trained on approximate values. [10]

As shown in the figure below, RF randomly chooses a feature set from the dataset and generates 'n' random trees from each set of attributes. The resultant output from all DTs was combined to form the final output. Algorithm of Random Forest is a statistical machine algorithm that uses multiple learning approaches to enhance the predictive accuracy. [5]

The two most crucial parameters while developing an optimized RF model are mtry, which denotes the number of randomly chosen features examined at each split, and n-tree, which refers to the total number of trees in the model. The Random Forest Classifier is the classifier ensemble which generates a number of decision trees from the subsets of the training data set and then combines the votes of various trees to generate the final output. [6]



Random forest architecture

RF works on the labeled dataset in order to make predictions and create a model, which is then used for classification of unlabeled data. The model makes use of the bagging methodology with random feature selection for making DTs in order to minimize the variance. The Random Forest Classifier is the classifier ensemble which generates a number of decision trees from the subsets of the training data set and then combines the votes of various trees to generate the final output.

Some of the merits associated with RF include that it is capable of assessing the significance of the variables in both regression and classification analysis, where the importance of the variables is measured based on the reduction of impurity at each node utilized for

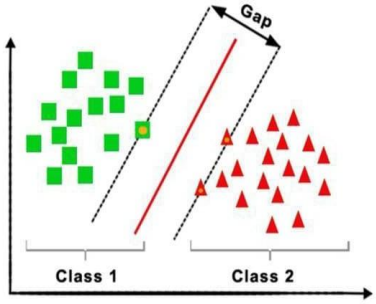
data partitioning. Moreover, it is able to handle the problem of missing values in the datasets automatically and mitigates the problem of overfitting prevalent with DTs. Moreover, it is well equipped to handle large datasets. However, there are also some demerits of RF, as it needs heavy computing power to give out results and requires a lot of training.

Support Vector Machine

Supervised learning methods include one of the most common approaches used in regression analysis and classification is known as support vector machines (SVMs). As seen in the image below, SVM works as a linear model to give solutions to linear as well as nonlinear cases. The concept of margins, which involves segmentation of data into individual categories and finding out correlations between them, is its core principle. SVM technique has been extensively used in medicine, biology, bioinformatics, and various other sciences. The role of SVM is that of building a better model that is being used in various applications. SVM uses the technique of building a hyperplane surface or multiple hyperplanes to train the data, which can be used for classification and regression purposes. [10]

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, October 2026
Impact Factor: 8.2

The support vector machine algorithm relies on statistical learning theory and adopts the principle of structural risk minimization. The aim of this method is to establish decision boundaries known as hyperplanes that would be able to distinguish the various classes by establishing the right hyperplane in an operational N-dimensional space. The decision boundary between the two classes is given by SVM by the assessment of the data points, especially those on the boundary line separating the two classes known as support vectors.



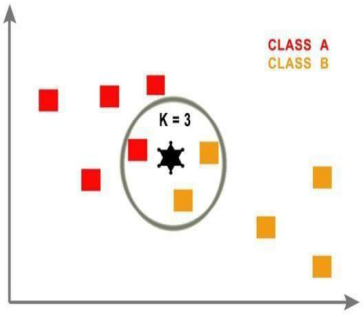
Support vector machine

SVM offers numerous advantages, such as its effectiveness with both semi-structured and unstructured data. [10] A significant strength of SVM is the kernel trick, which allows it to solve complex issues when coupled with proper capabilities, as well as manage high-dimensional data. [10] Furthermore, SVMs have a lesser risk of overfitting. However, SVM has some limitations. The training time for the model increases dramatically with larger datasets, and choosing the suitable kernel function might be difficult. In addition, SVM does not perform well in noisy cases. Some tools to implement SVM are SVM Light with C, LibSVM with Python, MATLAB, KERNlab, Scikit-Learn, Weka.

K-nearest neighbor

The K-nearest neighbor classifier is described as an "instance-based learning" algorithm. It is largely used to address the classification issues. KNN determines the closeness of the class labels of the nearest training data to that of the test data based on a specified K value in order to determine the class of the test data, as shown in the figure below. The algorithm then counts the nearest data points based on this K value and assigns a label to the new test data class. Typically, KNN determines the number of nearest training data points by setting K according to the formula: k = n^ (1/2), where n represents dataset size.

There are different strengths of KNN; for example, it is highly effective when the number of training data points is large. In addition, it is easy to use and extendable, supporting many qualities and distances. Moreover, it is able to work with multi-class data. Nevertheless, there are several weaknesses of KNN: it is hard to find appropriate K and hard to determine the suitable distance measure for a particular dataset. The other problem is that there is quite high cost of processing, because it requires calculations between each pair of training data points. The KNN algorithm is one of the important Supervised Machine Learning Algorithms, which can be used both for regression as well as classification tasks.[7]



K-nearest neighbor

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, October 2026
Impact Factor: 8.2

Naive Bayes

The Naive Bayes (NB) classifier revolves around the probabilistic foundation provided by the Bayes theorem. The implementation of Naive Bayes classifier is simple since it does not involve any complicated recursive parameter estimations. Thus, this classifier can be effectively used to handle very large datasets. Naive Bayes algorithm determines the probability of belonging to a class based on a given class label. It goes through the dataset in one pass only. The NB classifier functions on the principle that the appearance of a particular feature in a class is not related to the appearance of any other feature. NB is a classification method where the concept of independence and relation between all the attributes is defined as follows: It defines that the condition of a particular attribute in a class does not influence the condition of any other attribute in the same class. [12]

Naive Bayes has several advantages, including its ease of implementation, ability to produce excellent results with limited training data, compatibility with both continuous and discrete data, and success in handling multi-class predictions. Additionally, irrelevant traits have no substantial impact on its predictions. However, NB has certain limitations: the assumption is made that all the characteristics are independent, which might not always be true, it might run into problems with zero frequency, and its predictions are not always accurate. WEKA, Python, RStudio, and Mahout are examples of commonly used Naive Bayes implementation tools.

Comparison of different supervised machine learning methods are as follows:-

Linear regression

Advantages: -

Linear regression models are easy to understand and can forecast, classify, and predict

Training linear regression models is fast on large datasets and Eliminating overfitting by regularization

Disadvantages: -

Linear regression models require linearity between independent and dependent variables.

Logistic regression

Advantages:-

Excellent performance with small datasets and its output is interpretable as probability

Disadvantages:-

It only offers linear solutions and Compliant data assumptions are required

Decision trees

Advantages:-

•They manage categorical characteristics and perform well with large feature-count datasets

Disadvantages:-

• Interpretability of the ensemble model may be problematic

Random forest

Advantages:-

• Resistance to overfitting: the random forest is highly generalizable and scalable for larger datasets

Disadvantages: -

Random forest training is computationally expensive for large datasets and can be sensitive to hyper parameters

Support vector machine

Advantages: -

High-dimensional space for input and few irrelevant features

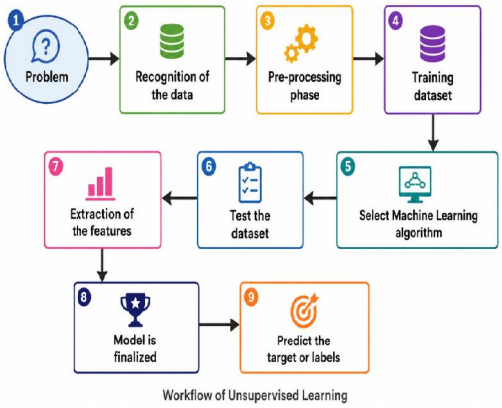
IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, October 2026
Impact Factor: 8.2

Disadvantages: -
•Data collection is time-consuming

K-nearest neighbors
Advantages: -
Simple, versatile, and non-parametric method
Disadvantages: -
• High computational complexity relative to other algorithms

Naive Bayes
Advantages:-
• Simple and direct technique that provides good accuracy
Disadvantages: -
Used primarily when the size of the training set is smaller

Unsupervised Learning: -
In contrast to supervised learning, unsupervised learning operates without definitive answers or instructors. This method is based on the premise that a machine can autonomously comprehend intricate processes and patterns without external assistance. It proves particularly advantageous in scenarios where experts are uncertain about what to search for within the data, and where the data itself lacks explicit objectives. The machine derives predictions from prior experiences and learns to forecast real-valued outcomes based on previously supplied information. Some well-known examples of unsupervised learning techniques include the clustering of data points by employing a similarity measure and dimensionality reduction to map high-dimensional data into lower-dimensional subspaces or feature selection. [1]



Workflow of Unsupervised Learning

Unsupervised learning is extensively applied in multimedia content processing, as the clustering and partitioning of data often necessitate the absence of class labels. K-means clustering, principal component analysis (PCA), and examples of unsupervised machine learning algorithms are the Apriori algorithm.

K-means
The k-means algorithm is a popular partitioning strategy and one of the most well-known unsupervised learning methods for solving the clustering problem. This technique divides a dataset into a preset number of clusters. K-means has various advantages, including higher computing efficiency than hierarchical clustering, especially when dealing with huge datasets. It also produces more compact clusters than hierarchical approaches when a lesser number of k is used. Furthermore, its simple execution and ease of comprehending the data are significant advantages. However, k-

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, October 2026
Impact Factor: 8.2

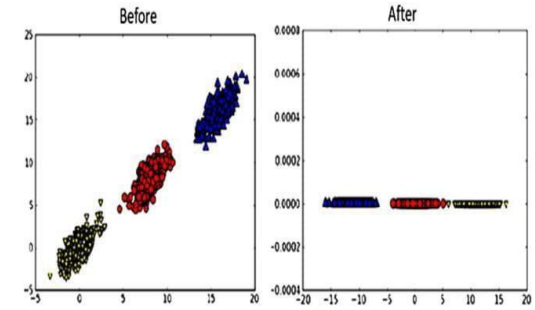
means has some downsides, including the difficulty of calculating the ideal value of K. Furthermore, variable initial conditions can result in varied final outcomes, which affect total performance.

Principal Component Analysis

Principal component analysis (PCA) is an important tool in modern data analysis since it allows critical information to be extracted from a dataset more easily. It successfully compresses data by keeping only the most important properties, reducing information loss and streamlining the dataset description.

PCA is frequently used to minimize the dimensionality of data before implementing classification models. Unsupervised approaches, methods like dimensionality reduction, clustering, and so forth may also be employed in visualizing data, recognizing overall trends or patterns, reducing the amount of data. PCA transforms two-dimensional data into one- dimensional data by converting the original set of variables into new variables known as principal components (PC), which are orthogonal to one another. This dimensionality reduction enhances the speed and ease of calculations.

To represent the functioning of PCA in the context of two-dimensional data. Through the application of PCA, the data is condensed into one dimension, as shown in the figure.



Visualization of data before and after applying PCA [49]

Apriori

Apriori algorithm uses iterative method called planar search using k element groups to explore (k + 1) element groups. It first creates a collection of frequent 1-element group by scanning the dataset to find the frequency of each element, and then selects elements that satisfy minimum support requirement. Such collection is called L1. L1 is then utilized to derive L2, and a recursive process continues to generate L3, and so forth, until there are no more frequent groups of k elements that can be found. For each Lk, scanning through the entire data set is necessary.

For the efficiency of generation of the elements for repeated occurrence in each level, the Apriori property is used, which decreases the search space. The Apriori property says that the subsets of the frequent elements set are also frequent. Two-step method, consisting of joins and prunes operations, is implemented to find groupings of similar elements.

Comparison of different unsupervised machine learning methods are as follows: -

K-means

Advantages:-

• Easily understandable algorithm with linear relative complexity

Disadvantages:-

User needs to mention the number of clusters or classes, and poor performance due to irregular shape cluster.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, October 2026
Impact Factor: 8.2

Principal component analysis

Advantages:-

• PCA helped reduce the complexity involved in image clustering and dataset size representation

Disadvantages: -

It is difficult to evaluate the covariance matrix accurately

Apriori

Advantages:-

Makes use of the large item set property and Simple to implement, parallelize

Disadvantages: -

• More scans of the database are needed and the cost of I/O operations increases because of these scans on the database of transactions

Reinforcement Learning: -

Reinforcement learning (RL) differs from both supervised learning and unsupervised learning in that reinforcement learning is goal-directed. In RL, an agent, which acts as a controller, is tasked with navigating the learning process to achieve specific objectives. The agent selects actions that affect the state of the environment, resulting in rewards. These incentives might be expressed as positive or negative numerical values. The main objective of the agent is to ensure maximum possible rewards from the environment in the long run. The field of RL has great prospects for changing many health care applications, and it has been used lately for checking context-aware symptoms for diagnosing diseases. [1] A task can be defined as the complete description of the environment, where the rewards are produced. The two commonly used algorithms in the reinforcement learning process include the Q-learning algorithm and the Monte Carlo tree search (MCTS).

Q-learning

It is a model-free RL method known as Q-Learning, which could be termed asynchronous dynamic programming. With Q-Learning, an agent can learn the best course of action in a Markov environment by looking at the effect of its own actions without requiring any mapping of the domain. Such a method is a progressive technique for dynamic programming which requires minimal computing power. It works by constantly refining the assessment of the quality of individual activities in certain situations.

Q-learning is frequently used in information theory, and further research is being conducted to explore its applications. Q-learning has recently been combined with information theory and employed in a variety of domains, including natural language processing (NLP), pattern identification, anomaly detection, and visual classification. In addition, a framework has been designed to generate suitable replies depending on user input in voice interaction systems utilizing reinforcement learning. Furthermore, a high-resolution deep learning-based prediction system for local rainfall has been established.

One of the primary advantages of developmental Q-learning is its capacity to accurately identify reward values in a multi-agent environment, where agents interact with one another. However, the challenge associated with Q-learning is that the output may end up being trapped in local minima since agents always strive to make the shortest possible paths.

Monte Carlo Tree Search

Monte Carlo Tree Search (MCTS) is an effective method to solve sequential decision-making problems. It uses an advanced tree search in order to effectively maintain exploration and exploitation. MCTS is based on random sampling in simulations as well as maintaining information about activity. This decision-making technique is very effective for navigating big, complex tree-like structures. In this context, each node denotes a state, often known as a problem configuration, and the edges indicate transitions between states.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 4, October 2026 Impact Factor: 8.2

MCTS is applicable to situations which can be represented using the Markov Decision Process (MDP), a form of discrete time stochastic control process. Certain adaptations of MCTS enable its application to partially observable Markov Decision Processes (POMDP). Recently, the combination of MCTS with deep reinforcement learning has formed the foundation of AlphaGo, developed by Google DeepMind. The fundamental MCTS technique is conceptually straightforward.

Comparison of different reinforcement machine learning methods are as follows:-

Q-learning

Advantages:-

It has broad applicability and can be used in solving problems which cannot be solved using any other techniques.
Implementation is easy and it can be applied in a number of different settings.

Disadvantages:-

• It may be hard to learn and may be highly sensitive to its initialization, in which case it will not converge to an optimal policy.

Monte Carlo tree search

Advantages:-

• It is helpful when dealing with a problem having a large number of search space.

Disadvantages: -

• It can be sensitive to its initial conditions, such that if it is not properly initialized, it will learn a less-than-optimal strategy.

Result and Discussion

This research carried out a systematic review to examine the recent advancements in machine learning (ML) and deep learning (DL) in healthcare predictions. This focus on healthcare predictions and the significance and dependability of ML and DL applications. This review mainly focuses on different ML and DL models which are used to predict in healthcare sector. These include the number of studies done for each of these parameters, methodology and maximum level of accuracy attained. However, it is important to consider the size of datasets employed in these training and validation process, as those trained on more extensive datasets are more likely to generalize well to novel data. Overall, the data show that ML and DL models are capable of making accurate predictions. Therefore, while choosing a model for particular application, it is important to choose from the various models and methodologies available.

Challenges

Although there have been great advancements in machine learning (ML) and deep learning (DL) in recent years, yet many barriers lie ahead in the path of applying these methods to resolve the major issues of the healthcare organizations.

In this regard, the following section focuses on some of the barriers that arise when trying to implement the machine learning and deep learning methods for predictions in healthcare. The first barrier is the Healthcare Data Stream. The healthcare industry is producing an enormous amount of medical data at a very fast pace, which is causing rapid changes in the industry. Many barriers arise with respect to analyzing this fast-moving and huge volume of streaming data.

Moreover, complexity in the healthcare domain constitutes a great barrier to machine learning and deep learning application. Healthcare research can sometimes be much more complex than other disciplines. There is an immense gap of information regarding the causes, transmission, and treatment of many diseases. Gathering enough data becomes particularly challenging due to the small number of patients that can be examined. But there are solutions that might be considered. The scarcity of patients requires profiling of the patients, techniques, innovations, and the process of

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, October 2026
Impact Factor: 8.2

integration of additional datasets. Specialists are able to study every dataset individually with suitable methods of machine learning before combining everything into one model.

II. CONCLUSION

In this case, the researcher examines the studies that have been carried out on the use of Machine Learning in health care. The Decision Tree and Support Vector Machine methods are commonly applied in predictive healthcare research since they are more accurate.

Machine Learning and Artificial Intelligence allow for almost unlimited opportunities in healthcare. In order to effectively manage health problems, they can assist in optimizing daily activities, disease assessment and prediction, planning of the treatment, and personalization of the healthcare services. More sophisticated Machine Learning methods are likely to appear in the future. They will be related to early disease detection, diagnosis, and prognosis. This research study shows that AI and Machine Learning can be applied for translation of health data into health benefits.

Machine Learning and Deep Learning can revolutionize the application of traditional healthcare delivery system through their application in prediction. Under such applications, healthcare data becomes an important component affecting the medical care systems. This paper aims to give a comprehensive analysis of the most important Machine Learning and Deep Learning technologies applied in the healthcare predictive analytics.[11]

The findings from the studies reviewed show that the use of AI techniques is crucial for proper diagnosis of ailments and analysis of data in health care. The research carried out contributes to the literature of predictive healthcare analytics through the application of Machine Learning and Deep Learning techniques.

REFERENCES

[1] J. Wiens and E. S. Shenoy, "Machine Learning for Healthcare: On the Verge of a Major Shift in Healthcare Epidemiology," *Clin. Infect. Dis.*, vol. 66, no. 1, pp. 149–153, Jan. 2018, doi: 10.1093/cid/cix731.

[2] M. Badawy, N. Ramadan, and H. A. Hefny, "Healthcare predictive analytics using machine learning and deep learning techniques: a survey," *J. Electr. Syst. Inf. Technol.*, vol. 10, no. 1, Aug. 2023, doi: 10.1186/s43067-023-00108-y.

[3] B. A. Mousa and B. Al-Khateeb, "Predicting medicine demand using deep learning techniques: A review," Jan. 01, 2023, *De Gruyter Open Ltd*. doi: 10.1515/jisys-2022-0297.

[4] A. Qayyum, J. Qadir, M. Bilal, and A. Al-Fuqaha, "Secure and Robust Machine Learning for Healthcare: A Survey," 2021, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/RBME.2020.3013489.

[5] D. Maulud and A. M. Abdulazeez, "A Review on Linear Regression Comprehensive in Machine Learning," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 2, pp. 140–147, Dec. 2020, doi: 10.38094/jastt1457.

[6] S. Rong and Z. Bao-Wen, "The research of regression model in machine learning field," in *MATEC Web of Conferences*, EDP Sciences, Jul. 2018. doi: 10.1051/matecconf/201817601033.

[7] S. Rahman, F. M. Javed, M. Shamrat, Z. Tasnim, J. Roy, and S. A. Hossain, "A Comparative Study On Liver Disease Prediction Using Supervised Machine Learning Algorithms," *Int. J. Sci. Technol. Res.*, vol. 8, no. 11, 2019, [Online]. Available: www.ijstr.org

[8] M. Manjiri, M. Mastoli, U. R. Pol, and R. D. Patil, "Machine Learning Classification Algorithms for Predictive Analysis in Healthcare," *Int. Res. J. Eng. Technol.*, 2019, [Online]. Available: www.irjet.net

[9] S. Kharya, "Using Data Mining Techniques for Diagnosis and Prognosis of Cancer Disease," *Int. J. Comput. Sci. Eng. Inf. Technol.*, vol. 2, no. 2, pp. 55–66, Apr. 2012, doi: 10.5121/ijcseit.2012.2206.

[10] V. Thanigaivasan, S. J. Narayanan, and N. C. Sriman Narayana Iyengar, "Analysis of Parallel SVM Based Classification Technique on Healthcare using Big Data Management in Cloud Storage," *Recent Patents Comput. Sci.*, vol. 11, no. 3, pp. 169–178, Aug. 2018, doi:10.2174/2213275911666180830145249.

[11] S. Moturi, "Classification Model for Prediction of Heart Disease using Correlation Coefficient Technique," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 2, pp. 2116-– 2123, Apr. 2020, doi: 10.30534/ijatcse/2020/185922020.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, October 2026
Impact Factor: 8.2

[12] M. O. Edeh *et al.*, "A Classification Algorithm-Based Hybrid Diabetes Prediction Model," *Front. Public Heal.*, vol. 10, Mar. 2022, doi: 10.3389/fpubh.2022.829519.

[13] K. Qu, "Research on linear regression algorithm," MATEC Web Conf., vol. 395, p. 01046, 2024, doi: 10.1051/matecconf/202439501046.

[14] D. Maulud and A. M. Abdulazeez, "A Review on Linear Regression Comprehensive in Machine Learning," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 2, pp. 140–147, Dec. 2020, doi: 10.38094/jastt1457.

[15] B. A. Mousa and B. Al-Khateeb, "Predicting medicine demand using deep learning techniques: A review," Jan. 01, 2023, *De Gruyter Open Ltd*. doi: 10.1515/jisys-2022-0297.

[16] S. Rong and Z. Bao-Wen, "The research of regression model in machine learning field," in *MATEC Web of Conferences*, EDP Sciences, Jul. 2018. doi: 10.1051/matecconf/201817601033.