Comparative Analysis of Machine Learning and Deep Learning Algorithms Used in Heart Disease Detection

K.Ramesh
Kakatiya Government College, Hanumakonda

Abstract- The early diagnosis of heart disease is a crucial area of medical study, and the application of machine learning (ML) and deep learning (DL) techniques in this field has yielded encouraging results. The identification of cardiac problems using ML and DL approaches has been the subject of earlier research, which is presented in this study along with a thorough survey and comparative analysis. A broad variety of supervised learning algorithms, such as decision trees (DT), support vector machines (SVM), random forest (RF), naive bayes (NB), k-nearest neighbours (KNN) as well as deep learning (DL) methods like convolutional neural networks (CNN), artificial neural networks (ANN), recurrent neural networks (RNN) and long short term memory (LSTM), are explored in this study. Their advantages and disadvantages are examined. The paper also analyses the difficulties in detecting cardiac illness, including data accuracy, interpretability, generalizability, and possible biases, and offers insights into potential future research directions.

Keywords- Heart Disease Detection, Machine Learning, Deep Learning, Supervised Learning Algorithms, Deep Neural Network

I. INTRODUCTION

Heart disease is one of the leading causes for mortality in many nations and is a widespread worldwide health problem. Patient care is significantly impacted by the capacity to effectively forecast and diagnose cardiac disease since it allows for early intervention, individualised treatment strategies, and better results. Machine learning (ML) and deep learning (DL) approaches have recently revolutionised the fields of healthcare and medical research by providing strong tools for analysing complicated medical data and making wise therapeutic judgements.

Around the world, heart illnesses have become one of the leading causes of mortality. 17.7 million people worldwide die each year from heart-related disorders, or 31% of all fatalities, according to the World Health Organisation. Heart-related illnesses are now the top cause of death worldwide, including in India [1]. Indians lost 1.7 million people to heart disease in 2016, according to the 2016 Global Burden of Disease Report, which was published on September 15, 2017. Heart-related illnesses raise healthcare costs and lower an individual's productivity. According to World Health Organisation (WHO) estimates, heart-related and cardiovascular disorders cost India $237 billion between 2005 and 2015 [2]. Hence, it is crucial to make reliable predictions of heart-related disorders.

It is difficult to anticipate heart disorders since they are influenced by several risk factors, genetic markers, lifestyle variables, and medical history. The complex interplay between these factors are sometimes difficult to capture using traditional rule-based methods. In contrast, substantial patient datasets, such as medical records, imaging data, genetic information, and lifestyle variables, may be analysed using machine learning and deep learning algorithms to spot minor patterns and correlations that can be a sign of a risk for heart disease. These frameworks are capable of generalising information from prior patient data to forecast the possibility of cardiac problems in new patients, allowing for early identification, individualised risk assessment, and specialised therapeutic techniques. These frameworks have the potential to increase prediction accuracy through ongoing training with new data, support healthcare professionals in making informed decisions, and ultimately improve patient outcomes and the management of heart diseases.

A thorough analysis of the various supervised ML and DL techniques used for heart disease prediction is presented in this survey. This research aims to obtain insights into the state-of-the-art methodologies, evaluate their effectiveness, and identify possible difficulties and future research initiatives in this field by performing an exhaustive literature study and critical analysis of existing studies [3]. Fig 1 provides an overview of different types of heart diseases that have increased fatality rates over the years. Table I summarizes the disease shown in Fig 1.

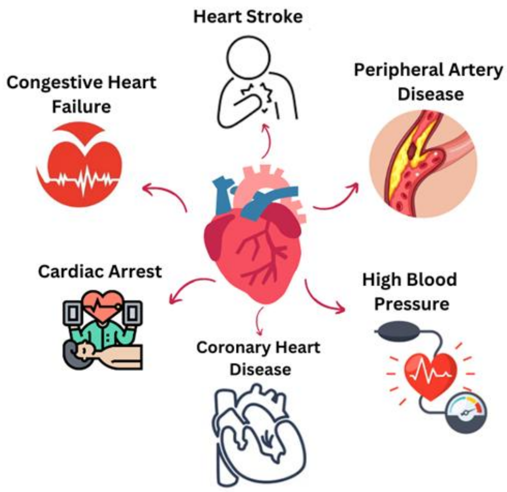


Fig 1. Overview of Heart Diseases

Table I. Different Types of Heart Diseases

Disease	Overview
Heart Stroke	Brain injury resulting from blood supply interruption.
Peripheral Artery Disease	The circulatory disorder is characterised by constricted blood arteries that restrict blood flow to the limbs.
High Blood Pressure	Condition where blood pushes too hard against the arterial walls.
Coronary Heart Disease	The main blood arteries of the heart are susceptible to injury or illness.
Cardiac Arrest	Heartbeat, consciousness, and breathing suddenly stop without warning.
Congestive Heart Failure	The chronic condition when the heart does not pump blood.

II. PAPER OBJECTIVE

The objectives of this survey are summarized as follows

1. To provide a thorough analysis of previous research papers and publications on heart disease prediction, with an emphasis on how machine learning and deep learning techniques were used.
2. To present the classification and evaluation of various heart disease prediction frameworks, such as decision trees (DT), support vector machines (SVM), random forest (RF), naïve bayes (NB), k nearest neighbours (knn), and deep neural network frameworks like convoluted neural network (CNN). artificial neural network (ANN), recurrent neural network (RNN), long short term memory (LSTM)..
3. To determine the benefits and drawbacks of the aforementioned methods used to forecast heart disease, including problems with the accuracy of the data, their interpretability, generalizability, and biases that could be present.
4. To identify areas in need of more study and innovation by highlighting emerging trends and research directions in cardiac disease prediction utilising ML and DL technology.

III. LITERATURE REVIEW

Numerous studies have been conducted employing various deep learning and machine learning algorithms to develop illness prediction systems.

Papers [4], [5] and [6] explore various deep learning techniques like convoluted neural networks (CNN), artificial neural networks (ANN), etc to detect cardiovascular diseases. Paper [4] uses optimized DNN using Talos hyperparameter. [7] uses machine learning techniques on the UCI dataset [8] to analyse the performance of the supervised learning algorithms. A non-linear classification system to predict cardiac disease has been put forward by [9]. For the prediction of heart disease using an optimised attribute set, bigdata techniques like Hadoop Distributed File System (HDFS), MapReduce, and SVM are suggested. This study investigated the application of several data mining approaches for the prognosis of cardiac disorders. It advises utilising HDFS to store vast amounts of data across several nodes and running the SVM-based prediction algorithm across multiple nodes at once. When SVM is utilised in parallel, the calculation time is faster than when SVM is used sequentially.

Prediction and Analysis of Heart Disease Occurrence Using Data Mining Techniques is advised by [10]. The major goal is to foresee the development of cardiac disease in order to quickly and automatically diagnose the condition. The suggested technique is crucial in healthcare organisations with professionals that lack current knowledge and expertise. To determine if a person has heart disease or not, many medical characteristics are used, such as blood sugar and heart rate, age, and sex. Dataset analyses are calculated using the WEKA programme.

An IoT and deep learning based framework is integrated in [11]. It proposes a system that acquires data from IoT devices, and the electronic clinical data stored on the cloud pertaining to patient history are subjected to predictive analytics. The smart healthcare system for monitoring and accurately predicting heart disease risk in this framework is built around Bi-LSTM (bidirectional long short-term memory).

A cardiovascular disease analysis is presented in [12]. This study suggested data mining methods for illness prediction. It is intended to offer an overview of the most recent methods for extracting data from datasets, which will be helpful to healthcare professionals. In [13] c45 rules and partial tree techniques are explored to predict heart disease. This paper can discover set of rules to predict the risk levels of patients based on given parameter about their health. The performance can be calculated in measures of accuracy classification, error classification, rules generated and the results.

Table II presents an overview of previous research works, their contributions and limitations of the methods used in them.

Table II: Overview of Previous Research Works

Paper	Year	Contribution	Method Used	Parameter Considered	Results - Accuracy Metrics	Limitations
[4]	2020	Overview of performance of deep learning methods	CNN, ANN, DNN	Talos Hyperparameter - cholesterol level, fasting blood sugar, electrocardiographic results, maximum heart rate - multivariate features.	CNN: 95.37%, ANN: 92.16%	Overlook of benefits of combining several models.
[5]	2020	Analysis of ANN in medical domain(cardio)	ANN	Features such as age, sex, chest pain type, resting blood pressure, cholesterol level, fasting blood sugar, electrocardiographic results.	ANN: 87.80%	Lack of exploration of drawbacks of the framework.
[6]	2020	Analysis of neural network frameworks that can be used to predict heart diseases	RNN, CNN	Attributes like cholesterol level, fasting blood sugar, electrocardiographic results derived from preprocessing and feature engineering.	CNN: 88.62%, RNN: 84.52%	Absence of in-depth analysis of the frameworks.
[7]	2021	Exploration of ML algorithms on UCI dataset	SVM, DT. RF	Parameters for the machine learning algorithms (e.g., kernel type, maximum depth for decision trees).	DT: 84.35%, SVM: 76.91% , RF: 88.95%	No detailed analysis of the hyperparameter tuning process.
[9]	2022	Hybrid model approach to detect heart disease	HDFS, MapReduce, SVM	Features such as age, sex, chest pain type, resting blood pressure, cholesterol level, fasting blood sugar, electrocardiographic results.	SVM: 86.24%	Lack of insights on the computational overhead while implementing the proposed system.

[10]	2022	Analysis of the occurrence of heart disease using data mining techniques.	NB, SVM	Smoothing parameter for NB.	SVM: 86%, NB: 73.29%,	Lack of analysis of potential dataset biases.
[11]	2022	Use of IoT and DNN for heart disease prediction	Bi-LSTM	Hyperparameters for the Bi-LSTM.	LSTM: 87.61%, Bi-LSTM: 93%	Scarcity of real world deployment analysis
[12]	2022	Use of data mining techniques	Feature Extraction Methods	Features such as age, sex, chest pain type, resting blood pressure, cholesterol level, fasting blood sugar, electrocardiographic results.		No proper comparison of other data mining techniques
[13]	2023	Heart Disease Prediction System Evaluation using C4.5	NB, C4.5	C4.5 decision tree algorithm - chest pain type, resting blood pressure, cholesterol level, fasting blood sugar, electrocardiographic results.	NB: 83%	Lack of discussion of solutions to the raised challenges.

IV. PROPOSED METHODOLOGY

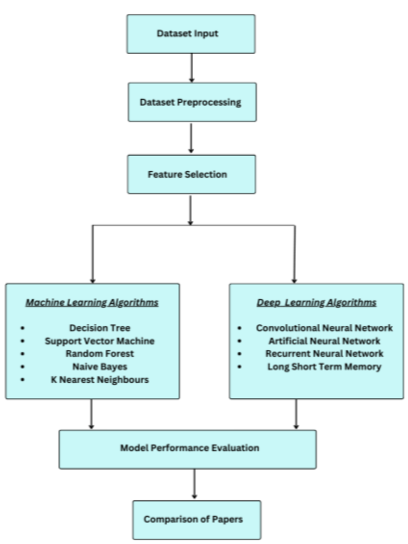


Fig 2. Proposed Methodology Flow

This review compares several ML and DL methodologies. Heart disease data with patient attributes are the first step in the process, which involves feature selection and data preparation. Deciphering is done for ML techniques like RF, DT, SVM, KNN, NB, and DL architectures like ANN, CNN, RNN, and LSTM. The end product is a thorough overview study that highlights the findings, benefits, and drawbacks of several ML and DL algorithms used to the prediction of heart disease. Fig. 2 summarises the proposed methodology workflow of this research initiative. The purpose of this survey paper is to offer beneficial insights to academics and professionals working in the area of healthcare and medical data analysis.

V. MACHINE LEARNING ALGORITHMS

This section provides an overview of the supervised machine learning algorithms that have been used over time for the detection of various heart diseases.

1. Decision Tree

Decision Tree is a popular machine learning technique noted for its interpretability and efficiency in a variety of applications, including the diagnosis of cardiac problems. To create predictions, it recursively divides the data into subsets according to the most important qualities, resulting in a tree-like structure.

Decision trees can identify patients as having a higher or lower risk of developing heart disease by examining patient data, including age, blood pressure, cholesterol levels, and other pertinent factors. The technique is well-suited for medical datasets with a variety of attributes since it can handle both numerical and categorical data [14].

While decision trees may predict cardiac illness with high accuracy, processing complicated and noisy data can be difficult for them, which might impair their ability to generalise [15]. Decision trees may be a useful tool in assisting with the early detection and risk assessment of cardiac illnesses due to ensemble approaches like Random Forests and Gradient Boosting that can be used to improve forecast accuracy and robustness. Table III presents the performance of the decision tree algorithm on the research works analysed in this survey.

Table III. Performance of DT on Heart Disease Prediction

Paper	Dataset Used	Accuracy
[7]	UCI Heart Disease Dataset	84.35%
[14]	Heart Failure Dataset	95.61%
[15]	Cardiovascular Disease Dataset	90%

2. Support Vector Machine

Support Vector Machine is a potent supervised machine learning method that is frequently employed for classification problems, such as the identification of heart disease. SVM separates distinct groups effectively by locating an ideal hyperplane that maximises the margin between various classes of data points. SVM performs admirably when used to forecast cardiac disease since it can handle both linearly and non-linearly separable data. SVM can accurately classify patients into various risk groups by revealing complex patterns and correlations in the data by translating the input characteristics into a higher-dimensional space.

SVM is particularly useful in medical applications because it can handle high-dimensional data and resist overfitting. In these applications, datasets may contain few samples but many features [16]. SVM has demonstrated promising results in the detection of heart disease, achieving high levels of accuracy, sensitivity, and specificity in differentiating between healthy people and those at risk of heart disease, thus providing invaluable support to clinicians in making informed decisions for prompt interventions and enhanced patient care [17]. Table IV presents the performance of the support vector machine algorithm on the research works analysed in this survey.

Table IV. Performance of SVM on Heart Disease Prediction

Paper	Dataset Used	Accuracy
[7]	UCI Heart Disease Dataset	76.91%
[9]	UCI Heart Disease Dataset	86.24%
[10]	UCI Heart Disease Dataset	86%
[16]	UCI Heart Disease Dataset	73%
[17]	Raw heart related medical data	92.30%

3. Random Forest

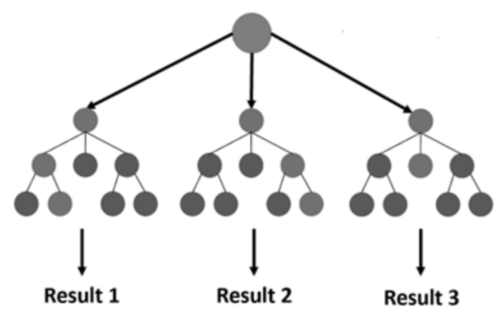


Fig 3. Random Forest Classifier Diagram

Random Forest, an ensemble model, has been popular in many fields, including medical diagnostics like the identification of heart problems. This method, which consists of numerous decision trees, integrates their results to provide reliable forecasts. Random Forest makes use of its capacity to manage high-dimensional data and identify intricate correlations between characteristics, which are typical in medical datasets [18], in the context of heart disease diagnosis. It is especially well suited for healthcare applications with sparse and unbalanced data because to its capacity to minimise overfitting and reduce bias. Fig. 3 shows how random forest combines base level models, combining them to obtain the final result.

Additionally, the interpretability of Random Forest makes it possible for doctors to comprehend the model's decision-making process, fostering trust and making it easier to incorporate AI-driven solutions into clinical practise. Due of these benefits, Random Forest has shown promising results in the identification of cardiac disease, obtaining high accuracy [19]. Table V presents the performance of the random forest algorithm on the research works analysed in this survey.

Table V. Performance of RF on Heart Disease Prediction

Paper	Dataset Used	Accuracy
[7]	UCI Heart Disease Dataset	88.95%
[18]	Kaggle Cardio Dataset	75.44%
[19]	UCI Heart Disease Dataset	90.81%

4. Naïve Bayes

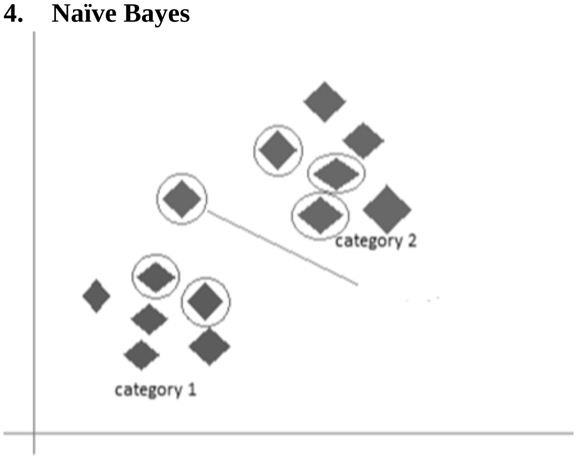


Fig 4. Naïve Bayes Classifier Diagram

Naïve Bayes is based on the Bayes theorem and makes the "naive" assumption that characteristics are conditionally independent given the class label. It can perform remarkably well in heart disease identification despite its naive assumptions since it can handle high-dimensional data and operate with very small datasets that are frequently found in medical applications. Fig. 4 shows the graph where two categories are grouped based on the plotted data and new data point is highlighted from the clusters.

Real-time diagnosis and risk stratification are two applications where the method excels due to its computational efficiency and quick training period. NB is also capable of dealing with discrete and continuous characteristics, which makes it possible to accommodate a variety of medical data, including patient demographics, test findings, and EKG readings [20]. NB is still an important tool in the medical field, especially when computational resources and data availability are constrained, even if it may not attain the same degree of prediction accuracy as more sophisticated models like deep neural networks [21]. Its simplicity also makes it simple to read, making it easier for doctors to comprehend the model's judgements, which is important for promoting trust and acceptance in clinical settings [22]. Table VI presents the performance of the naive bayes algorithm on the research works analysed in this survey.

Table VI. Performance of NB on Heart Disease Prediction

Paper	Dataset Used	Accuracy
[10]	UCI Heart Disease Dataset	73.29%
[13]	UCI Heart Disease Dataset	83%
[20]	UCI Heart Disease Dataset	96.99%
[21]	UCI Heart Disease Dataset	89.24%
[22]	UCI Heart Disease Dataset	82.25%

5. K-Nearest Neighbours

KNN is a machine learning technique that is utilised for classification tasks, such as the diagnosis of cardiac disorders. Similar cases are likely to have the same class label, which is the foundation of KNN. By taking into account the class labels of its k-nearest neighbours in the feature space, KNN identifies the class of an unseen data point in the context of heart disease prediction.

Despite its ease of use, KNN has demonstrated encouraging results in the detection of cardiac illnesses, particularly when used with well-structured and well-annotated datasets. It has the benefit of being non-parametric, meaning it doesn't make assumptions about the underlying data distribution, and it can successfully capture non-linear correlations between features [23].

However, the right selection of k, the quantity of neighbours, and the calibre of the data are vital for its effectiveness, making rigorous preprocessing and feature selection essential. Additionally, KNN's computational cost in high-dimensional feature spaces might be a drawback, affecting its performance on bigger datasets [24]. Despite this, KNN can be a trustworthy and understandable choice when customised and optimised for the specific heart disease prediction job, offering helpful insights into the early diagnosis and risk assessment of cardiovascular disorders. Table VII presents the performance of the k nearest neighbours algorithm on the research works analysed in this survey.

Table VII. Performance of KNN on Heart Disease Prediction

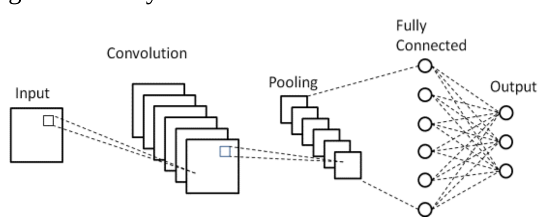
Paper	Dataset Used	Accuracy
[23]	Cardiovascular Dataset	73.28%
[24]	UCI Heart Disease Dataset	89.50%

VI. DEEP LEARNING FRAMEWORKS

This section explores the deep learning frameworks experimented on for the detection of heart diseases.

1. Convoluted Neural Network

Fig. 5 CNN Layers



In the area of medical image analysis, particularly the diagnosis of cardiac problems, CNNs have become a potent deep learning method. CNNs can effectively extract complicated patterns and features from detailed medical pictures, allowing them to identify minute characteristics that may be essential for a precise diagnosis. CNNs have performed remarkably well in the context of detecting heart illness by efficiently analysing angiograms, cardiac MRI scans, echocardiograms, and other imaging modalities [25]. CNNs are able to recognise anatomical structures, detect abnormalities, and forecast a variety of cardiac diseases with high sensitivity and specificity by taking use of their capacity to automatically learn hierarchical representations. Fig. 5 shows the different layers of a CNN framework where classification is followed by feature extraction layers, Additionally, CNNs may be customised and adjusted for various cardiac disease classification jobs, enabling the risk categorization and treatment planning of individual patients.

Output sizes of a CNN can be calculate using:

$$n_{out} = [(n_{in} + 2p - k)/s] + 1$$

Where,

n_{out} = number of output features

n_{in} = number of input features

p = convolution padding size

k = convolution kernel size

s = convolution stride size

CNNs have substantially enhanced cardiac imaging and have shown tremendous potential in improving early identification and diagnosis of heart illnesses, ultimately leading to better patient outcomes and more effective healthcare practices [26]. Table VIII presents the performance of the CNN algorithm on the research works analysed in this survey.

Table VIII. Performance of CNN on Heart Disease Prediction

Paper	Dataset Used	Accuracy
[4]	Kaggle Heart Dataset	95.37%
[6]	UCI Heart Disease Dataset	88.62%
[25]	Coronary Artery Image Set	95.74%
[26]	Heart Failure Prediction Dataset (Kaggle)	93%

2. Artificial Neural Network

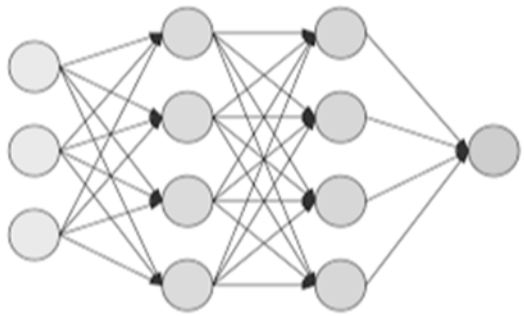


Fig. 6 ANN Layers

Due to its capacity to recognise intricate patterns and connections in vast medical datasets, ANN has become an effective tool for the early identification of cardiac disease. The neuronal structure of the human brain served as the inspiration for the class of deep learning models known as ANNs, which are composed of linked layers of artificial neurons that analyse and change incoming data. When used to forecast heart disease, ANNs perform very well, demonstrating excellent sensitivity and accuracy in identifying diverse cardiovascular diseases. Fig. 6 shows the hidden layers of an artificial neural network internal framework.

ANNs can find hidden patterns and risk factors linked to cardiac ailments by utilising their ability to automatically extract characteristics from raw data, enabling early diagnosis and individualised treatment strategies [27]. Additionally, ANNs' flexibility in handling various medical data types, such as electrocardiograms (ECGs), medical imaging, and patient demographics, further improves their accuracy in identifying and categorising cases of cardiac disease. But the availability of high-quality and diverse datasets, as well as rigorous model tweaking and validation to prevent overfitting, are crucial for ANNs' effectiveness in the identification of cardiac disease. Over time, ANN-based strategies have shown to be an effective weapon in the battle against heart illnesses, providing hope for better cardiac treatment and preserving lives.

3. Recurrent Neural Network

RNNs are ideally suited for applications in the detection of cardiac disorders since they have become potent deep learning models for sequential data processing. RNNs are able to analyse sequences of varying lengths and recognise patterns across time since they are built to capture temporal relationships in data. RNNs can efficiently handle time-series data obtained from electrocardiograms (ECGs) or heart rate variability signals in the context of heart disease diagnosis [28]. RNNs are able to recognise minute variations and enduring patterns suggestive of cardiac abnormalities, such as arrhythmias or aberrant heart rhythms, by taking advantage of the sequential nature of these physiological signals.

Due to its capacity to acquire temporal dynamics and contextual information, RNNs have shown promise in the detection of heart disorders [29]. This is because they can precisely discriminate between anomalies linked to various cardiovascular ailments and normal cardiac rhythms. To enable the reliable and broad deployment of RNN-based cardiac disease detection systems in clinical practise, it is imperative to solve issues related to dataset size, model interpretability, and generalisation [30].

4. Long Short Term Memory

Due to its capacity to recognise long-range relationships and temporal patterns, LSTM is a RNN that has demonstrated to be very successful in a number of sequential data tasks. With its ability to store knowledge over long periods of time, LSTM models have shown outstanding efficacy in the context of heart disease diagnosis [31]. Electrocardiograms (ECGs) and vital signs are examples of sequential physiological data that may be analysed to identify minor patterns suggestive of heart problems and identify potential cardiac dangers [32]. The capacity of LSTM to understand complicated associations and handle time-series data allows it to identify irregular heartbeats, arrhythmias, and other cardiac defects with remarkable precision [33].

The gates of LSTM have the following equation:

$$i_t = \sigma(w_i[\boldsymbol{h}_{t-1}, w_t] + b_i$$

Where,

i_t = input gate

$\boldsymbol{\sigma}$ = sigmoid function

w_i = weight of ith gate

h_{t-1} = output of the previous LSTM gate

x_t = input at current timestamp

bi = biases of the ith gate

The field of medicine has taken a keen interest in it because to its efficacy in detecting heart illness, underlining LSTM's potential as a useful tool in aiding healthcare practitioners with early diagnosis and efficient risk stratification for cardiovascular patients [34].

VII. ANALYSIS

ML and DL have both proved their potential ability over years of research. Choice of suitable frameworks depends upon a number of variables, including the size of the dataset, the complexity of the features, and the availability of computer resources. Deep learning, with its capacity to automatically deduce complex patterns and hierarchies from unstructured data, has demonstrated promising outcomes in a number of medical applications, including the prediction of heart disease. It is especially helpful in identifying subtle patterns that conventional machine learning techniques would find difficult to identify because of its ability to handle unstructured data, such as time-series data and medical imaging. Deep learning, however, frequently necessitates a substantial quantity of labelled data and computationally costly training, which might be a drawback in the context of medical imaging.

For leveraging predictions, Methods for extracting features from raw data are essential for converting them into understandable representations. Researchers can considerably enhance the effectiveness of conventional machine learning algorithms for heart disease prediction by carefully choosing and developing essential characteristics [35]. Medical data may be enhanced by extracting important features using feature selection methods, dimensionality reduction techniques, and domain-specific expertise, making it more suitable for the selected machine learning model.

VIII. CHALLENGES

There are various obstacles that call for careful attention when using machine learning and deep learning techniques to identify cardiac disease. First, the efficacy of models depends on the correctness of the data used for training and testing them. Medical data may be unreliable, lacking, or unbalanced, which might cause biases in the learnt models. Additionally, black-box algorithms may make it difficult for doctors to comprehend and believe the predictions made by DL and complicated ML models, raising serious concerns about their interpretability in the medical field. Given the differences in data distributions and patient demographics, ensuring the generalizability of cardiac disease detection models across various patient groups and healthcare settings is another difficulty [36].

Furthermore, racial or gender biases that are present in the data itself may be carried over into the models, resulting in differences in predictions and significant discrepancies in healthcare outcomes. In order to produce trustworthy and equitable DL and ML solutions for heart disease diagnosis, it is necessary to address these problems with a careful approach to data quality, model interpretability, rigorous assessment methodologies, and ethical concerns [37].

IX. FUTURE SCOPE

By obtaining a more thorough picture of a patient's health, the integration of multi-modal data, such as genomics, wearable sensor data, and electronic health records, might increase the predictive potential of models. Additionally, investigating the potential of domain adaptation and transfer learning techniques may help models generalise across various healthcare contexts and demographics, overcoming the difficulties posed by data heterogeneity. The development of innovative explainable DL and ML models, the use of attention processes, and a focus on interpretable AI research might increase the transparency of prediction models and promote adoption in clinical practise.

To guarantee fair healthcare delivery, it is also crucial to look at the ethical ramifications of using AI-driven heart disease prediction models, including possible biases and their influence on health inequalities [38]. The seamless translation of cutting-edge research into clinical applications will be made possible by encouraging collaboration among AI researchers, clinicians, and healthcare stakeholders. This will ultimately improve patient outcomes and raise the general effectiveness of cardiac healthcare systems.

X. CONCLUSION

This research survey has offered a thorough overview of ML and DL methods for heart disease prediction. It presents a critical assessment of the cutting-edge approaches used in this field by an extensive study of the literature and critical analysis of the already published research. Deep neural networks demonstrated promising performance in the detection of cardiac illnesses by efficiently capturing temporal patterns and long-range relationships, as emphasized by the evaluation of several ML and DL techniques, which also revealed their strengths and limits. To ensure the accuracy, fairness, and generalizability of predictive models, a number of issues including data correctness, interpretability, generalizability, and possible biases must be resolved. Interdisciplinary partnerships between academics, doctors, and data scientists are crucial to advancing cardiac disease prediction because they help to develop models, enhance data quality, and resolve ethical issues.

REFERENCE

[1] Ramadoss and Shah B et al.“A. Responding to the threat of chronic diseases in India”. Lancet. 2005; 366:1744–1749. doi: 10.1016/S0140-6736(05)67343-6.

[2] Global Atlas on Cardiovascular Disease Prevention and Control. Geneva, Switzerland: World Health Organization, 2011.

[3] Marimuthu, M., M. Abinaya, K. S. Hariesh, K. Madhankumar, and V. Pavithra. "A review on heart disease prediction using machine learning and data analytics approach." *International Journal of Computer Applications* 181, no. 18 (2018): 20-25.

[4] Sharma, Sumit, and Mahesh Parmar. "Heart diseases prediction using deep learning neural network model." *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* 9.3 (2020): 2244-2248.

[5] Pasha, S. N., Ramesh, D., Mohmmad, S., & Harshavardhan, A. (2020, December). Cardiovascular disease prediction using deep learning techniques. In *IOP conference series: materials science and engineering* (Vol. 981, No. 2, p. 022006). IOP Publishing.

[6] Ali, Farman, Shaker El-Sappagh, SM Riazul Islam, Daehan Kwak, Amjad Ali, Muhammad Imran, and Kyung-Sup Kwak. "A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion." *Information Fusion* 63 (2020): 208-222.

[7] Boukhatem, Chaimaa, Heba Yahia Youssef, and Ali Bou Nassif. "Heart disease prediction using machine learning." *2022 Advances in Science and Engineering Technology International Conferences (ASET)*. IEEE, 2021.

[8] Detrano, R., Janosi, A., Steinbrunn, W., Pfisterer, M., Schmid, J., Sandhu, S., Guppy, K., Lee, S., & Froelicher, V. (1989). International application of a new probability algorithm for the diagnosis of coronary artery disease. American Journal of Cardiology, 64,304--310.

[9] R. Sharmila, S. Chellammal, “A conceptual method to enhance the prediction of heart diseases using the data techniques”, International Journal of Computer Science and Engineering, May 2022.

[10] Mr. Chala Beyene, Prof. Pooja Kamat, “Survey on Prediction and Analysis the Occurrence of Heart Disease Using Data Mining Techniques”, International Journal of Pure and Applied Mathematics, 2022.

[11] Nancy, A. Angel, et al. "Iot-cloud-based smart healthcare monitoring system for heart disease prediction via deep learning." *Electronics* 11.15 (2022): 2292.

[12] Sharan Monica.L, Sathees Kumar.B, “Analysis of CardioVasular Disease Prediction using Data Mining Techniques”, International Journal of Modern Computer Science, vol.4, 1 February 2022, pp.55-58.

[13] Sharma Purushottam, Dr. Kanak Saxena, Richa Sharma, “Heart Disease Prediction System Evaluation using C4.5 Rules and Partial Tree”, Springer, Computational Intelligence in Data Mining,vol.2, 2023, pp.285-294

[14] Reddy, V. Sai Krishna, et al. "Prediction on Cardiovascular disease using Decision tree and Naïve Bayes classifiers." *Journal of Physics: Conference Series*. Vol. 2161. No. 1. IOP Publishing, 2022.

[15] El Massari, Hakim, Noreddine Gherabi, Sajida Mhammedi, Zineb Sabouri, and Hamza Ghandi. "Ontology-based decision tree model for prediction of cardiovascular disease." *Indian J. Comput. Sci. Eng* 13, no. 3 (2022): 851-859.

[16] Nugroho, K. S., Sukmadewa, A. Y., Vidianto, A., & Mahmudy, W. F. (2022). Effective predictive modelling for coronary artery diseases using support vector machine. *IAES International Journal of Artificial Intelligence*, *11*(1), 345.

[17] Nayak, Subrata Kumar, Sateesh Kumar Pradhan, Sujogya Mishra, Sipali Pradhan, and P. K. Pattnaik. "Prediction of Cardiac Arrest Using Support Vector Machine and Rough Set." In *2022 9th International*

Conference on Computing for Sustainable Global Development (INDIACom), pp. 164-172. IEEE, 2022.

[18] El-Shafiey, Mohamed G., et al. "A hybrid GA and PSO optimized approach for heart-disease prediction based on random forest." *Multimedia Tools and Applications* 81.13 (2022): 18155-18179.

[19] Wang, J., Rao, C., Goh, M. and Xiao, X., 2023. Risk assessment of coronary heart disease based on cloud-random forest. *Artificial Intelligence Review*, *56*(1), pp.203-232.

[20] Sameer, S. K. L., and P. Sriramya. "Improving the accuracy for prediction of heart disease by novel feature selection scheme using decision tree comparing with naive-bayes classifier algorithms." In *2022 International Conference on Business Analytics for Technology and Security (ICBATS)*, pp. 1-8. IEEE, 2022.

[21] Radpour, Vafa, and Farhad Soleimanian Gharehchopogh. "A Novel Hybrid Binary Farmland Fertility Algorithm with Naïve Bayes for Diagnosis of Heart Disease." *Sakarya University Journal of Computer and Information Sciences* 5, no. 1 (2022): 90-103.

[22] Gbadamosi, Babatunde, Roseline Oluwaseun Ogundokun, Emmanuel Abidemi Adeniyi, Sanjay Misra, and Nkiruka Francisca Stephens. "Medical Data Analysis for IoT-Based Datasets in the Cloud Using Naïve Bayes Classifier for Prediction of Heart Disease." In *New Frontiers in Cloud Computing and Internet of Things*, pp. 365-386. Cham: Springer International Publishing, 2022.

[23] Prabha, R., M. Razmah, T. Veeramakali, S. Sridevi, and R. Yashini. "Machine Learning Heart Disease Prediction Using KNN and RTC Algorithm." In *2022 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS)*, pp. 1-5. IEEE, 2022.

[24] Thummala, G.S.R. and Baskar, R., 2022, July. Prediction of Heart Disease using Decision Tree in Comparison with KNN to Improve Accuracy. In *2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)* (pp. 1-5). IEEE.

[25] Jain, Arushi, Annavarapu Chandra Sekhara Rao, Praphula Kumar Jain, and Yu-Chen Hu. "Optimized levy flight model for heart disease prediction using CNN framework in big data application." *Expert Systems with Applications* 223 (2023): 119859.

[26] Rustam, F., Ishaq, A., Munir, K., Almutairi, M., Aslam, N. and Ashraf, I., 2022. Incorporating CNN Features for Optimizing Performance of Ensemble Classifier for Cardiovascular Disease Prediction. *Diagnostics*, *12*(6), p.1474.

[27] Ravish, D. K., K. J. Shanthi, Nayana R. Shenoy, and S. Nisargh. "Heart function monitoring, prediction and prevention of heart attacks: Using artificial neural networks." In *2014 International Conference on Contemporary Computing and Informatics (IC3I)*, pp. 1-6. IEEE, 2014.

[28] Krishnan, Surenthiran, Pritheega Magalingam, and Roslina Ibrahim. "Hybrid deep learning model using recurrent neural network and gated recurrent unit for heart disease prediction." *International Journal of Electrical & Computer Engineering (2088-8708)* 11, no. 6 (2021).

[29] Pillai, N.S.R., Bee, K.K. and Kiruthika, J., 2019. Prediction of heart disease using rnn algorithm. *International Research Journal of Engineering and Technology*, *5*.

[30] Javid, Irfan, Ahmed Khalaf Zager Alsaedi, and Rozaida Ghazali. "Enhanced accuracy of heart disease prediction using machine learning and recurrent neural networks ensemble majority method." *International Journal of Advanced Computer Science and Applications* 11, no. 3 (2020).

[31] Maragatham, G., and Shobana Devi. "LSTM model for prediction of heart failure in big data." *Journal of medical systems* 43 (2019): 1-13.

[32] Djerioui, Mohamed, Youcef Brik, Mohamed Ladjal, and Bilal Attallah. "Heart Disease prediction using MLP and LSTM models." In *2020 International Conference on Electrical Engineering (ICEE)*, pp. 1-5. IEEE, 2020.

[33] Junwei, Kuang, Hangzhou Yang, Liu Junjiang, and Yan Zhijun. "Dynamic prediction of cardiovascular disease using improved LSTM." *International Journal of Crowd Science* 3, no. 1 (2019): 14-25.

[34] Dileep, P., Kunjam Nageswara Rao, Prajna Bodapati, Sitaratnam Gokuruboyina, Revathy Peddi, Amit Grover, and Anu Sheetal. "An automatic heart disease prediction using cluster-based bi-directional LSTM (C-BiLSTM) algorithm." *Neural Computing and Applications* 35, no. 10 (2023): 7253-7266.

[35] Shah, S. M. S., Batool, S., Khan, I., Ashraf, M. U., Abbas, S. H., & Hussain, S. A. (2017). Feature extraction through parallel probabilistic principal

component analysis for heart disease diagnosis. *Physica A: Statistical Mechanics and its Applications*, *482*, 796-807.

[36] OMKARI, D. Y., & SHINDE, S. B. (2023). OPPORTUNITIES AND CHALLENGES OF MACHINE LEARNING AND DEEP LEARNING TECHNIQUES IN CARDIOVASCULAR DISEASE PREDICTION: A SYSTEMATIC REVIEW. *Journal of Biological Systems*, *31*(02), 309-344.

[37] Kresoja, Karl-Patrik, Matthias Unterhuber, Rolf Wachter, Holger Thiele, and Philipp Lurz. "A cardiologist's guide to machine learning in cardiovascular disease prognosis prediction." *Basic research in cardiology* 118, no. 1 (2023): 10.

[38] Wong, Dragon YL, Mary C. Lam, Anran Ran, and Carol Y. Cheung. "Artificial intelligence in retinal imaging for cardiovascular disease prediction: current trends and future directions." *Current Opinion in Ophthalmology* 33, no. 5 (2022): 440-446.