International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 8 Issue: 3
DOI: https://doi.org/10.17762/ijfrcsce.v8i3.2087
***Article Received**: 10 June 2022 **Revised**: 20 August 2022 **Accepted**: 25 August 2022 **Publication**: 15 September 2022*

A Comparison Analysis of Machine Learning Algorithms on Cardiovascular Disease Prediction

Lakkala Jayasree
[1]Research Scholar, Department of C.S.E.,
Dr. M.G.R Educational & Research Institute, Chennai, India.
email: jayasreemohan15@gmail.com

Dr. D. Usha
Associate Professor, Department of C.S.E,
Dr. M.G.R Educational & Research Institute, Chennai, India.
email: ushahits@gmail.com

Abstract—People nowadays are engrossed in their daily routines, concentrating on their jobs and other responsibilities while ignoring their health. Because of their hurried lifestyles and disregard for their health, the number of people becoming ill grows daily. Furthermore, most of the population suffers from a disease such as cardiovascular disease. Cardiovascular disease kills 35% of the world's population, according to W.H.O. A person's life can be saved if a heart disease diagnosis is made early enough. Still, it can also be lost if the diagnosis is constructed incorrectly. Therefore, predicting heart disease will become increasingly relevant in the medical sector. The volume of data collected by the medical industry or hospitals, on the other hand, can be overwhelming at times. Time-series forecasting and processing using machine learning algorithms can help healthcare practitioners become more efficient. In this study, we discussed heart disease and its risk factors and machine learning techniques and compared various heart disease prediction algorithms. Predicting and assessing heart problems is the goal of this research.

Keywords- cardiovascular issue, machine learning, heart attack prediction, algorithms, data mining

I. Introduction

Overemphasis is placed on healthcare systems rather than technology. Technology has dramatically aided many medical applications, such as diagnosis, risk assessment, and prognosis. Machine Learning (ML) is a technique used in computer-aided medical applications that assists doctors and caregivers in forecasting patient situations and tailoring treatment accordingly [1]. The primary goal of using ML in medicine is to provide patients with chronic diseases with reliable clinical procedures. Current data-driven and intelligent healthcare systems must be adapted to assist chronic disease patients and caregivers. Cardiovascular and respiratory diseases are the leading causes of death worldwide, according to the W.H.O. Cardiovascular diseases claim the lives of 17.5 million people annually.

In contrast, chronic respiratory disorders claim the lives of 1.59 million. The most frequent types of heart disorders are shown in Table 1. The insufficiency of various healthcare systems hampers disease cure. Patients' trust in medical treatments declines as a result. Both doctors and patients can profit from a quick response to a patient's urgent state. Researchers worldwide are attempting to develop intelligent and efficient medical healthcare systems that lower the chance of death from fatal diseases [2]. It is possible to predict a patient's current state using modern methods like Nave Bayes, Random Forests (R.F.), and Neural Networks (N.N.) [3]. These methods send out health alerts based on a patient's previous medical history. As far as we know, these procedures only anticipate a patient's current condition. They are all uninterested in forecasting a patient's future situation. A patient's disease in an acute medical setting can deteriorate in a split second—response times in an emergency range from 1 to 3 minutes. Doctors may be jeopardized due to a lack of information about their patients' illnesses. Forecasting a patient's future status is critical to effective therapy. Predicting a patient's future state may also save lives and provide helpful information to caregivers [4]. Predicting cardiac disease has been a significant focus of many researchers. The strategies used to forecast cardiac disease are thoroughly examined in this paper.

Table 1: Numerous categories of heart diseases

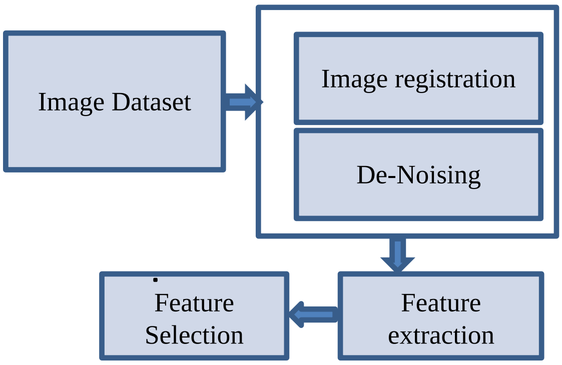
Condition	A person's heartbeat is out of whack
Congestive heart failure	Chronic illness impairs the heart's ability to pump blood efficiently.
Coronary artery disease	The heart's major blood arteries are damaged or diseased.
Peripheral artery disease	Circulatory disorder causes narrowed blood arteries in the limbs.
Cardiac arrest	Consciousness and respiration are abruptly interrupted.

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 8 Issue: 3
DOI: https://doi.org/10.17762/ijfrcsce.v8i3.2087
***Article Received**: 10 June 2022 **Revised**: 20 August 2022 **Accepted**: 25 August 2022 **Publication**: 15 September 2022*

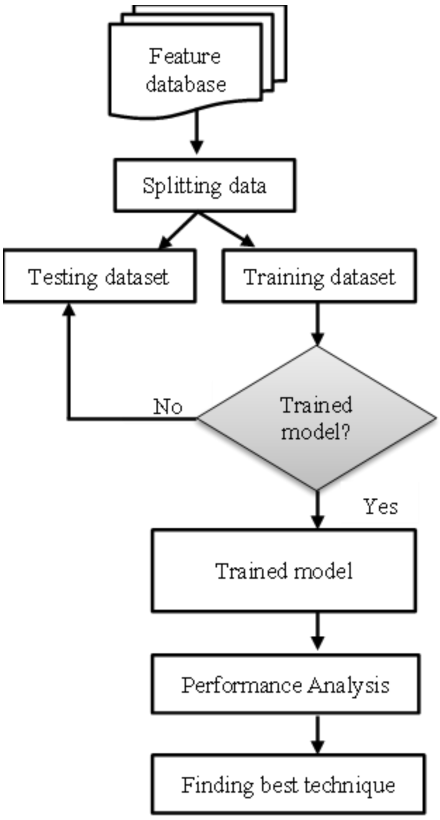
High Blood Pressure	Blood pressure on the artery walls is too high, causing it to malfunction.
Stroke	Damage to the brain occurs when the blood supply is cut off.

II. ML PARADIGM FOR HEART DISEASE PREDICTION

The conventional paradigm used by many researchers to predict cardiovascular disease is depicted in Figure 1. Two phases are involved: the first phase extracts the image's features. Predictions of illness are made using ML algorithms in phase 2.



(a) Phase 1- Feature Selection



(b) Phase 2 – Methodology for prediction

Figure 1: paradigm used by researchers to predict heart disease.

The pre-processing unit classifies image registration from medical raw image databases. This registration aligns the image for de-noising on a primary image database due to speckle noises in medical images. As a result, de-speckling medical images with a filtering technique enhances classification outcomes. The selection of variable subset attributes is used in modern design. A large number of retrieved features can be used to identify the sample. They don't get in the way of prediction or classification difficulties. Feature classification is the second stage of the model architecture. It is necessary to categorize the data that has been extracted. This classification is based on the content and size of the file—a subset of the input class-oriented collection. To better predict unknown data class items, classification is utilized to locate and classify them. Each class model's prediction will be calculated after separating data class objects. There are two types of datasets: training and testing. The datasets are trained to classify them using the feature database. Simultaneously, 30% of the datasets will be used to test the proposed model's speed and accuracy. The process can be iterated more accurately by implementing error minimization techniques in classifiers. For example, the most efficient ML models for medical image classification will be used here, and the N.B. is a simple probabilistic classifier. Another method is to use Support Vector Machines (SVM) to find patterns in a dataset [5].

Data sets

Current research commonly employs the U.C.I. dataset, an open-source repository for academics. Cardiovascular risk factors include age, cigarette use, blood cholesterol, exercise, blood pressure, and stress. A heart abnormality is usually discovered as the patient's disease progresses. The primary concerns were then sourced. Age, sex, blood pressure, and fatness were all significant risk factors in the study. The machine indicated if the patient was in cardiac arrest. The American Heart Association acquired data for 50 people. Most cardiac patients have many difficulties, and each problem is considered one attribute in the datasets.

III. COMPARATIVE ANALYSIS OF ML ALGORITHMS

Classic models, ensemble classifiers, neural networks, and deep learning methods are the commonly used models for prediction, yet each model has shortcomings. As an example (1). White-box models, while interpretable, are less accurate than black-box models. On the other hand, Black boxes are slow and difficult to read. (2). When classes are separable linearly, linear models such as L.R. and SVM function well, while neural networks and deep learning can handle non-linear scenarios. (3). specifying machine learning model parameters, For

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 8 Issue: 3
DOI: https://doi.org/10.17762/ijfrcsce.v8i3.2087
***Article Received**: 10 June 2022 **Revised**: 20 August 2022 **Accepted**: 25 August 2022 **Publication**: 15 September 2022*

example, the K-NN. The algorithm needs the K. The selection of parameters for some models, such as deep learning, is more complicated. It may also be challenging to determine the parameters. (4). overfitting is a problem in machine learning training. To avoid overfitting, D.T. and L.R. models can be penalized with variable penalty strength. Tables 2 and 3 highlight conventional machine learning techniques [6]. Some academics employ various datasets in their studies, such as Cleveland, Hungary, Switzerland, and Long Beach V.

Table 2: Merits of ML algorithms to various properties

Properties	Algorithm									
	D.T.	LR	KNN	SVM	NB	RF	GB	XGBoost	NN	DL
Interpretability	√	√	√		√					
Speed	√	√	√		√					
Overfitting control		√		√						
Few Parameters			√							
Easy implementation			√		√					
Applicable to high dimensional data					√				√	√
Usage of diverse data						√		√		
Tackle the missing data							√			
Applicability with categorical data					√		√			
Greater accuracy				√		√	√	√	√	√

Table 3: Demerits of ML algorithms to various characteristics

	Algorithms									
	DT	LR	KNN	SVM	NB	RF	GB	XGBoost	NN	DL
Cannot handle more unrelated features	√	√	√							
Least accuracy	√	√	√		√					
High Memory	√		√	√						
Less numeric variables		√						√		
Not suitable for high dimensional data			√							
Requires huge data									√	√
Need for Expertize to tune the data										√
Initializing the learning parameters									√	√

*D.T.- Decision Tree, L.R.- Linear Regression, K.N.N.- K-nearest neighbours, N.B.- Naïve Bayes, GB-Gradient Boost, NN-Neural Network, DL –Deep Learning

IV. LITERATURE REVIEW

A full review was undertaken of a variety of early detection methods employed by several researchers. Over the years, massive data mining techniques have been used to construct a model for predicting heart disease.

Many studies have used machine learning algorithms to classify and predict cardiac disease. The scientist's employed CART

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 8 Issue: 3
DOI: https://doi.org/10.17762/ijfrcsce.v8i3.2087
***Article Received**: 10 June 2022 **Revised**: 20 August 2022 **Accepted**: 25 August 2022 **Publication**: 15 September 2022*

(Classification and Regression) to identify patients at high and low risk of congestive heart failure [7]. Deep neural networks could select and apply the best features to improve performance, say the authors of [8]. Other machine learnings and deep learning techniques were compared to SVM, RBFs and CARTs. More than 87.6% of the data was correctly categorized thanks to Random Forest and CART. Unstructured clinical records were used by Zhang et al. [9] to calculate a heart disease prediction rate of 93.37 percent.

Data with a large dimensionality is a common issue in machine learning. Real-time datasets are enormous, and we frequently can't even view them in 3D. As a result, working with this data necessitates a lot of memory, and the data can increase, leading to overfitting. The inclusion of weighting characteristics decreases dataset redundancy, allowing for speedier processing. Modifications to the dataset's structure can reduce its dimensionality [10]. The use of features and engineering improves classification and prediction results. [11] Deep learning and machine learning were used to diagnose heart problems and improve hyper settings. The other accurate models were logistic regression, SVM, and Random Forest. A binary classifier with Fisher as the ranking mechanism was used to avoid overfitting and increase training speed.

Several scholars have offered dimensionality reduction methodologies such as principal component analysis for dealing with vast amounts of multidimensional data. The authors in [12] attained an F1 score of 99.83 percent using a neural network classifier and five unsupervised dimensionality reduction approaches. Zhang et al. [13] used AdaBoost, a PCA-based method, to accomplish their results. Heart disease is a severe medical condition that can lead to death if left untreated. For men, heart disease is a greater risk than for women, according to the Harvard Medical School (H.M.S.). [14]. There is a two-to-to-to ratio between male and female heart attack risk. High cholesterol is just one risk factor; others include high blood pressure, diabetes, a high B.M.I., and not exercising enough. Data from 1998 in the public health dataset is critical to the study since it may be used as a benchmark to predict heart disease. Table 4 summarises the literature review.

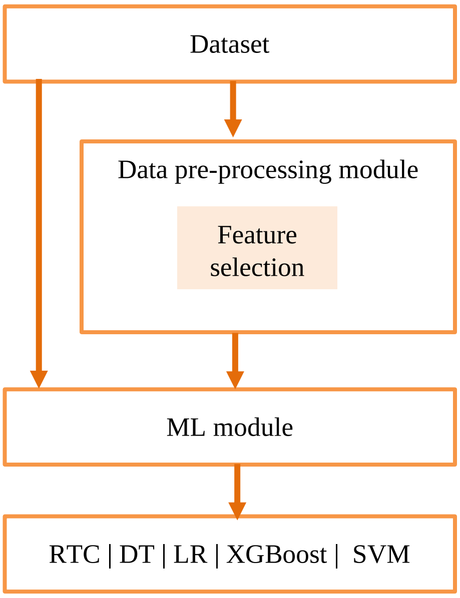
Table 4: Recent works on heart disease prediction

Author	Year	Findings
G´arate-Escamila et al [15]	2020	With the X2 statistical model, D.N.N. and ANN were used in conjunction.
H.M.S. [14]	2020	P.C.A. was used to reduce dimensionality and pick features from Hungarian-Cleveland datasets.
Senthilkumar Mohan et al [16]	2019	R.F. and Linear Method Characteristics Combination: (L.M.)
Liaqat Ali et al [17]	2019	Two models are used in this technique. Unlike SVM, the SVM model is linear and L1 regularised.
Zhang et al. [18]	2018	It was possible to improve the prediction accuracy by combining the AdaBoost classifier with P.C.A.
Singh et al. [19]	2018	Significant features were discovered using generalized discriminant analysis with binary classifiers.
Chen et al. [20]	2018	Subspace feature clustering was utilized to keep the number of distinct characteristics in each cluster to a manageable level.
Yang and Nataliani [21]	2018	Authors employed fuzzy clustering and weighted feature techniques to cut down on characteristics.
Kumar [22]	2017	Researchers used various machine learning techniques to arrive at these conclusions and then compared their findings.
Rajagopal and Ranganathan [23]	2017	A domain expert correctly analyzed the result using a probabilistic neural network classifier.
Zhang et al. [24]	2017	The New York Heart Association uses a set of codes to classify clinical data.

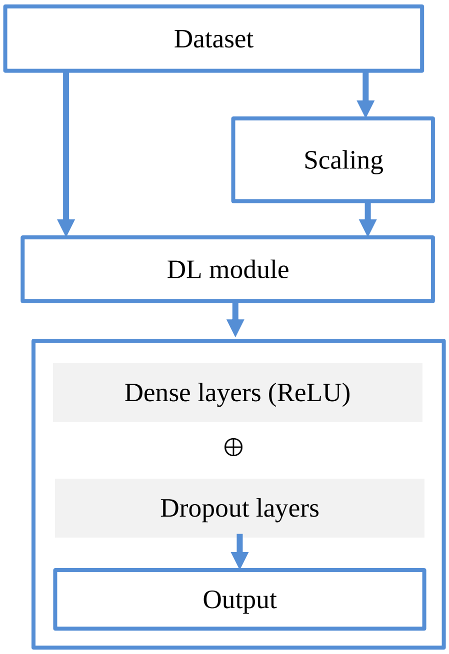
International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 8 Issue: 3
DOI: https://doi.org/10.17762/ijfrcsce.v8i3.2087
***Article Received**: 10 June 2022 **Revised**: 20 August 2022 **Accepted**: 25 August 2022 **Publication**: 15 September 2022*

Khan and Quadri [25]	2016	This study used unstructured data mining techniques to identify the optimal model and angiographic illness state.
Dun et al. [26]	2016	Much deep learning, ensemble techniques, and hyperparameter tuning were used to improve precision.

Other authors have recently conducted research in this area, evaluating various machine learning and deep learning algorithms for heart illness prediction. The authors of [27], for example, compare and contrast the results and analyses of the U.C.I. Machine Learning Heart Disease dataset uses various machine learning algorithms and deep learning methodologies. The methods used by the authors in their inquiry are depicted in Figure 2. The dataset has 14 essential attributes that are utilized to analyze the information. The accuracy and confusion matrix is then used to confirm several possible results. Non-relevant features are removed from the dataset using Isolation Forest, and the data is normalized for more reliable results. This research can be used in conjunction with other multimedia technologies, such as mobile devices, which are also discussed. Thanks to deep learning algorithms, the accuracy was 94.2 percent.



(a)



(b)

Figure 2: First and Second Schematic diagrams of the model

The authors in [28] use random Forest Bayesian Classification and Logistic Regression to predict heart disease in humans or those with risk factors. This model has 18 features and 1990 data points after pre-processing. The data was separated into 80 percent training and 20 percent testing sets. Using actual patient medical records, specific tests assessed the effectiveness and accuracy of the proposed system. Random forest accuracy of 92.44 percent, N.B. Classifier accuracy of 62%, and L.R. method accuracy of 59% are all promising findings for this system: its speed and ease of use.

Figure 3 depicts the implementation of the Random Forest, Decision Tree, and Hybrid Model machine learning algorithms [29]. FIGURE 3: (Hybrid of R.F. and D.T.). Experimental results show that the hybrid model's accuracy in predicting heart disease is 88.7 percent. The authors used a hybrid model of D.T. and R.F. approaches to gathering the user's input parameter for forecasting heart disease through the user interface.

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 8 Issue: 3
DOI: https://doi.org/10.17762/ijfrcsce.v8i3.2087
***Article Received**: 10 June 2022 **Revised**: 20 August 2022 **Accepted**: 25 August 2022 **Publication**: 15 September 2022*

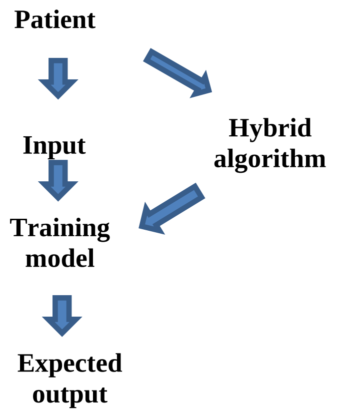


Figure 3:The hybrid strategy for prediction of disease

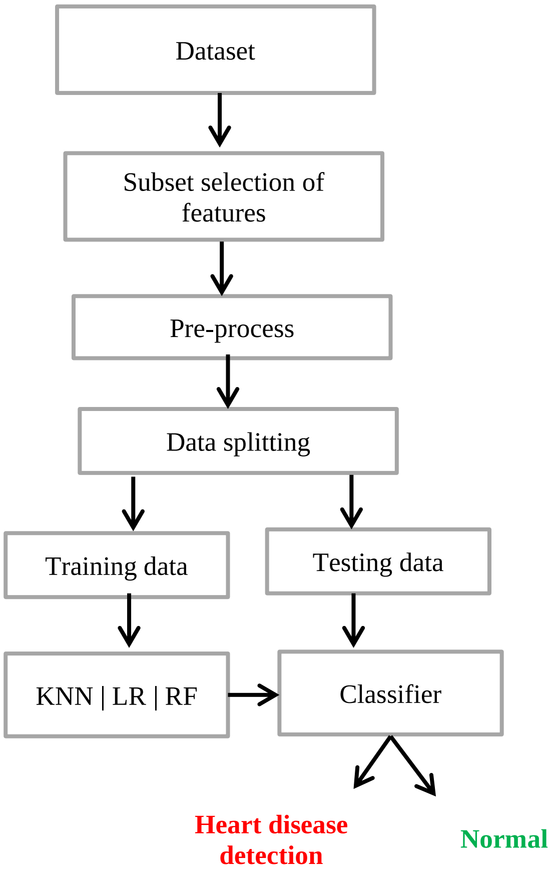


Figure 4: Model used for heart disease prediction [30]

According to the authors in [30], various classifiers have been utilized to generate an EHDPS. This model considers 13 medical factors to forecast the prediction, and the methodology is shown in Figure 4.

The efficiency of the algorithms can be measured using performance metrics. Accuracy is assessed using five different performance metrics. A specific feature can be calculated using only one performance metric's unique equation. The accuracy implied by Eq. (1), discussed further below, is used to evaluate the algorithm's efficiency. Eq. (2) mentions precision, a second measure that gives positive probability over the exactness of optimistic possibility. Eq. (3) depicts a sensitivity measure that can correctly detect the disease. The 'Specificity' measure, represented in Eq., can be used to determine the accurate detection of negative cases (4). The final standard, F-measure, can be used to evaluate the accuracy of a test and is denoted in Eq (5).

$$\text{Accuracy} = \frac{T_n+T_p}{T_n+T_p+F_n+F_p} \quad (1)$$

$$\text{Precision}= \frac{T_p}{T_p+F_p} \quad (2)$$

$$\text{Sensitivity}= \frac{T_p}{T_p+F_n} \quad (3)$$

$$\text{Specificity} = \frac{T_n}{T_n+F_p} \quad (4)$$

$$\text{F-measure}= \frac{2T_p}{2T_p+F_n+F_p} \quad (5)$$

The U.C.I. dataset is used in [30], with 13 attributes, and mainly focuses on detecting heart diseases effectively. Early detection of heart disease leads to a reduction in medical expenses. We can achieve this by applying different techniques and assessing their performance to get exact outputs. The used models are based on supervised learning paradigms like RF., NB, SVM, DT., and K-NN. The comparison results on the U.C.I dataset are shown in Table 5.

Table 5. Comparison of Accuracy

Method	Accuracy
R.F.	87%
NB	82%
SVM	65%
D.T.	73%
K.N.N.	62%

A study was conducted on several approaches and their performances using various paradigms. A projected model is proposed, which incorporates integrated elements and multiple tactics and classification algorithms.

Table 6. Various evaluation metrics on numerous models. [31]

Models	A	P	S	Sp	F
NB	76%	91%	80%	61%	84%
L.M.	86%	89%	95%	20%	92%
L.R.	83%	90%	91%	26%	90%
D.T.	86%	86%	99%	0%	92%

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 8 Issue: 3
DOI: https://doi.org/10.17762/ijfrcsce.v8i3.2087
***Article Received**: 10 June 2022 **Revised**: 20 August 2022 **Accepted**: 25 August 2022 **Publication**: 15 September 2022*

R.F.	87%	87%	99%	11%	93%
SVM	87%	86%	100%	0%	93%
HRFLM	88%	91%	93%	83%	91%

Several popular methods are examined, and the Hybrid Random Forest with a Linear Model (HRFLM) strategy was determined to have an accuracy level of 88.72 percent, as shown in Table 6. A, P, S, Sp. F corresponds to Accuracy, Precision, Sensitivity, Specificity, and F-Score, respectively. The authors compared the outcomes with available related theories. In [32], [33], [34] and [35] developed their systems with the Naïve Bayes paradigm and attained 82.05% accuracy with the use of 15 features dataset, 99.10% accuracy with 12 features dataset, 85% accuracy by using 15 features dataset and 97.03% accuracy by using seven features dataset respectively. We compared the models and found that the Naïve Bayes classifier performs well with higher precision with the dataset of fewer features than with many features. Table 7 compares the [36] model to the Naive Bayes Classifier with the other investigations.

Table 7: N.B. Classifier comparison of [36] model

Study	No. of features	Accuracy
[32]	15	83%
[33]	12	99%
[34]	15	86%
[35]	7	98%
[36]	19	63%

Furthermore, the model is compared to the R.F. model used to build the system [37] and [38]. They achieved an accuracy of 84.01 percent utilizing 15 datasets, while [38] employed 15 features to achieve a 77.03 percent accuracy before selecting a component. They used eight characteristics in the end and attained a 70.62 percent accuracy. The Random Forest approach performed well when the dataset had many features but poorly when the dataset had a small number of features. We compared the results of these models with the other recent studies in Table 8.

Table 8: R.F. algorithm comparison [36].

Study	No. of features	Accuracy
[37]	15	85%
[38]	15	78%
[38]	8	72%
[36]	19	94%

V. Conclusion

This study and subsequent research revealed that heart disease is a severe risk. Cardiovascular disease risk factors have been identified. The symptoms of various types of heart disease are also explained in this study. In this paper, we assess different machine learning approaches for predicting cardiac disease using a variety of benchmark datasets. According to a research review, only a tiny percentage of patients with heart disease can be expected. To better forecast early-onset heart illness, multi-modelling and sophisticated models are necessary. The scalability and accuracy of this prediction system could be improved in several ways. Deep learning is increasingly being utilized to boost performance in various fields. We intend to create other deep learning algorithms that outperform the Random Forest algorithm in the future.

References

[1]. Sengan, Sudhakar, et al. "Secured and privacy-based I.D.S. for healthcare systems on E-medical data using machine learning approach." *International Journal of Reliable and Quality E-Healthcare (IJRQEH)* 11.3 (2022): 1-11.

[2]. Fukuoka, Yoshimi, and Yoo Jung Oh. "Perceived Heart Attack Likelihood in Adults with a High Diabetes Risk." *Heart & Lung* 52 (2022): 42-47.

[3]. M. Tarawneh and O. Embarak, "Hybrid approach for heart disease prediction using data mining techniques," Advances in Internet, Data and Web Technologies, in Proceedings of the International Conference on Emerg- ing Internetworking, Data & Web Technologies, pp. 447–454, Springer, Fujairah Campus, U.A.E., February 2019

[4]. F. Jabeen, M. Maqsood, M. A. Ghazanfar et al., "An iot based efficient hybrid recommender system for cardiovascular disease," Peer-to-Peer Networking and Applications, vol. 12, no. 5, pp. 1263–1276, 2019.

[5]. Ali, Md Mamun, et al. "Heart disease prediction using supervised machine learning algorithms: performance analysis and comparison." *Computers in Biology and Medicine* 136 (2021): 104672.

[6]. Javan, Samaneh Layeghian, and Mohammad Mehdi Sepehri. "A predictive framework in healthcare: Case study on cardiac arrest prediction." *Artificial Intelligence in Medicine* 117 (2021): 102099.

[7]. P. Melillo, N. De Luca, M. Bracale, and L. Pecchia, "Classification tree for risk assessment in patients suffering from congestive heart failure via long-term heart rate variability," IEEE Journal of Biomedical and Health Informatics, vol. 17, no. 3, pp. 727–733, 2013.

[8]. M. M. A. Rahhal, Y. Bazi, H. Alhichri, N. Alajlan, F. Melgani, and R. R. Yager, "Deep learning approach for active classification of electrocardiogram signals," Information Sciences, vol. 345, pp. 340–354, 2016

[9]. R. Zhang, S. Ma, L. Shanahan, J. Munroe, S. Horn, and S. Speedie, "Automatic methods to extract New York heart association classification from clinical notes," in Proceedings of the 2017 IEEE International Conference on

Bioinformatics and Biomedicine (BIBM), pp. 1296–1299, IEEE, Kansas City, MO, U.S.A., November 2017
[10]. M.-S. Yang and Y. Nataliani, "A feature-reduction fuzzy clustering algorithm based on feature-weighted entropy," IEEE Transactions on Fuzzy Systems, vol. 26, no. 2, pp. 817– 835, 2018.
[11]. R. S. Singh, B. S. Saini, and R. K. Sunkaria, "Detection of coronary artery disease by reduced features and extreme learning machine," Medicine and Pharmacy Reports, vol. 91, no. 2, pp. 166–175, 2018
[12]. R. Rajagopal and V. Ranganathan, "Evaluation of effect of unsupervised dimensionality reduction techniques on automated arrhythmia classification," Biomedical Signal Processing and Control, vol. 34, pp. 1–8, 2017
[13]. D. Zhang, L. Zou, X. Zhou, and F. He, "Integrating feature selection and feature extraction methods with deep learning to predict clinical outcome of breast cancer," IEEE Access, vol. 6, pp. 28936–28944, 2018.
[14]. Harvard Medical School, "*roughout life, heart attacks are twice as common in men as women," 2020, https://www.health.harvard.edu/heart-health/throughout-life-heartattacks-are-twice-as-common-in-men-than-women
[15]. A. K. Garate-Escamila, A. Hajjam El Hassani, and E. Andr' es, ` "Classification models for heart disease prediction using feature selection and P.C.A.," Informatics in Medicine Unlocked, vol. 19, Article ID 100330, 2020.
[16]. Mohan, Senthilkumar, Chandrasegar Thirumalai, and Gautam Srivastava, "Effective heart disease prediction using hybrid machine learning techniques." IEEE Access 7 (2019): 81542-81554.
[17]. Ali, Liaqat, et al, "An optimized stacked support vector machines based expert system for the effective prediction of heart failure." IEEE Access 7 (2019): 54007-54014.
[18]. D. Zhang, L. Zou, X. Zhou, and F. He, "Integrating feature selection and feature extraction methods with deep learning to predict clinical outcome of breast cancer," IEEE Access, vol. 6, pp. 28936–28944, 2018
[19]. R. S. Singh, B. S. Saini, and R. K. Sunkaria, "Detection of coronary artery disease by reduced features and extreme learning machine," Medicine and Pharmacy Reports, vol. 91, no. 2, pp. 166–175, 2018.
[20]. R. Chen, N. Sun, X. Chen, M. Yang, and Q. Wu, "Supervised feature selection with a stratified feature weighting method," IEEE Access, vol. 6, pp. 15087–15098, 2018.
[21]. M.-S. Yang and Y. Nataliani, "A feature-reduction fuzzy clustering algorithm based on feature-weighted entropy," IEEE Transactions on Fuzzy Systems, vol. 26, no. 2, pp. 817– 835, 2018
[22]. S. Kumar, "Predicting and diagnosing of heart disease using machine learning algorithms," International Journal of Engineering and Computer Science, vol. 6, no. 6, pp. 2319–7242, 2017.
[23]. R. Rajagopal and V. Ranganathan, "Evaluation of effect of unsupervised dimensionality reduction techniques on automated arrhythmia classification," Biomedical Signal Processing and Control, vol. 34, pp. 1–8, 2017.
[24]. R. Zhang, S. Ma, L. Shanahan, J. Munroe, S. Horn, and S. Speedie, "Automatic methods to extract New York heart association classification from clinical notes," in Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1296–1299, IEEE, Kansas City, MO, U.S.A., November 2017
[25]. S. S. Khan and S. M. K. Quadri, "Prediction of angiographic disease status using rule based data mining techniques," Biological Forum—An International Journal, vol. 8, no. 2, pp. 103–107, 2016.
[26]. B. Dun, E. Wang, and S. Majumder, "Heart disease diagnosis on medical data using ensemble learning," 2016.
[27]. Bharti, Rohit, et al. "Prediction of heart disease using a combination of machine learning and deep learning." *Computational intelligence and neuroscience* 2021 (2021).
[28]. Obasi, Thankgod, and M. Omair Shafiq. "Towards comparing and using Machine Learning techniques for detecting and predicting Heart Attack and Diseases." *2019 IEEE international conference on big data (big data)*. IEEE, 2019.
[29]. Kavitha, M., et al. "Heart disease prediction using hybrid machine learning model." *2021 6th International Conference on Inventive Computation Technologies (ICICT)*. IEEE, 2021.
[30]. Jindal, Harshit, et al. "Heart disease prediction using machine learning algorithms." *I.O.P. Conference Series: Materials Science and Engineering*. Vol. 1022. No. 1. I.O.P. Publishing, 2021.
[31]. Kondababu, A., et al. "A comparative study on machine learning based heart disease prediction." *Materials Today: Proceedings* (2021).
[32]. S. Manikandan, "Heart attack prediction system," 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, 2017, pp. 817- 820
[33]. Hlaudi Daniel Masethe, Mosima Anna Masethe, "Prediction of Heart Disease using classification Algorithm", Proceedings of the World Congress on Engineering & Computer Science 2014, WCECS.
[34]. M. A. Jabbar and S. Samreen, "Heart disease prediction system based on hidden naïve bayes classifier," 2016 International Conference on Circuits, Controls, Communications and Computing (I4C), Bangalore, 2016, pp. 1-5.
[35]. Patel, S. B., Yadav, P. K. and Shukla, D. D., "Predict the diagnosis of heart disease patients using classification mining techniques", IOSR Journal of Agriculture and Veterinary Science, Vol. 4, No. 2, (2013), 61-64.
[36]. Obasi, Thankgod, and M. Omair Shafiq. "Towards comparing and using Machine Learning techniques for detecting and predicting Heart Attack and Diseases." *2019*

International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE)
ISSN: 2454-4248 Volume: 8 Issue: 3
DOI: https://doi.org/10.17762/ijfrcsce.v8i3.2087
***Article Received**: 10 June 2022 **Revised**: 20 August 2022 **Accepted**: 25 August 2022 **Publication**: 15 September 2022*

IEEE international conference on big data (big data). IEEE, 2019.

[37]. Kumar, Mukesh & Shambhu, Shankar & Sharma, Abha. (2018). Classification of Heart Diseases Patients using Data Mining Techniques. 10.13140/RG.2.2.11358.28486/1

[38]. Khateeb, Nida, Usman Muhammad "Efficient Heart Disease Prediction System using K-Nearest Neighbor Classification Technique" Proceedings of the International Conference on Big Data and Internet of Things, London, United Kingdom, December 20-22, 2017

.