ISSN 1816-6075 (Print), 1818-0523 (Online)
Journal of System and Management Sciences
Vol. 13 (2023) No. 6, pp. 40-60
DOI:10.33168/JSMS.2023.0603

Prediction of Heart Disease Using Machine Learning: A Systematic Literature Review

Alfredo Daza Vergaray[1], Juan Carlos Herrera Miranda[2], Juana Bobadilla Cornelio[3], Atilio Rubén López Carranza[4] and Carlos Fidel Ponce Sanchez[5]

[1] Faculty of Engineering and Architecture, School of Systems Engineering, Universidad César Vallejo, Lima, Peru

[2] Faculty of Systems Engineering, Professional Academic Career of Systems Engineering, Universidad Andina Néstor Cáceres Velásquez, Puno, Peru

[3] Faculty of Pedagogy and Physical Culture, School of Alternative Basic Education – Primary, Universidad Nacional de Educación Enrique Guzmán y Valle, Lima, Peru

[4] Faculty of Engineering, Civil Engineering Study Program, Universidad Privada Antenor Orrego, Trujillo, Peru

[5] Faculty of Industrial and Systems Engineering, School of Industrial Engineering, Universidad Nacional de Ingeniería, Lima, Peru

adaza@ucv.edu.pe (Corresponding author)

Abstract. This article aims to carry out a systematic review of the research works that deal with the topic of Machine learning (ML) and Deep learning (DL) to predict heart disease. For this purpose, an exhaustive search was carried out and after the analysis of the documentation collected, aspects such: countries have more studies been done on use of ML and DL, techniques most used and that had the best accuracy, tools, metrics, kind of heart disease and variable selection algorithms, which served as the basis for the elaboration of this document. The intent is to contribute to more profound understanding of methodologies, techniques and metrics with the applications of machine learning for predict heart disease. The results of the study showed that India, China and Pakistan were the countries with most studies on the use of ML and DL to predict heart disease, also Random Forest, SVM and Logistic Regression were the most used techniques, of which XGBoost, Ensemble Deep learning and Stacking were the ones that obtained the best accuracy results. Python was the tool considered the best. The most concurrent metrics used was Accuracy, Precision and F1-Score, the type of disease that has been applied was Coronary Artery Heart, as the selection algorithms were Kernel and Information Gain. Due to the paucity of this studies on heart disease using machine learning and deep learning, this work also points the way to new research.

Keywords: Machine Learning, Deep Learning, Heart disease, Prediction.

1. Introduction

Cardiovascular diseases (CVD) are one of the leading causes of death worldwide (Ahsan et al., 2021). Thus, it is estimated that in 2019, 17.9 million people died from CVD, representing this figure to 32% of all deaths worldwide, which 85% of deaths are due to strokes or strokes (World Health Organization, 2021); furthermore, statistics in Saudi Arabia collected over the past 40 years show that deaths have been increasing (Aljefree, Shatwan and Almoraie, 2021); Among the causes of heart disease, there are different problems, such as: dietary risks, particle pollution, high blood pressure, high body mass index, cholesterol, etc., as well as health habits and aging (Vaduganathan, et al., 2022; Hall and Hall, 2016).

However, it is important to detect people at high risk of CVD early, this to provide treatments that help prevent unexpected deaths, so about three-quarters of deaths from cardiovascular diseases (CVD) occur in low- and middle-income countries (World Health Organization, 2021). Therefore, with the high number of people with heart disease, it is increasingly difficult to provide necessary diagnoses to patients in less developed places such as: India, Africa and Bangladesh, whose processes of detecting these diseases and their symptoms are doubtful, this because there is still low accessibility of appropriate technological equipment (Lip et al., 2017); At the same time, the places of care that exist, it is unaffordable for the population and they fail to take advantage of the opportunities for an accurate diagnosis and cure plans (Angraal et al., 2019). Thus, techniques based on machine learning (ML) and Deep learning (DL), have been applied in different sectors, including: entertainment (Justesen et al., 2019; Berno et al., 2021), education (Daza et al., 2022; Daza et al., 20229, Technology, Commerce, as well as in health specializations (Ferreras et al., 2023; Mendo et al.,2021; Almeida and Tavares, 2020; Alonso et al., 2018; Ting et al., 2019), while if it is about predicting early risks in people with heart disease, it is done automatically based on clinical information (Bhowmick et al.,2022; Erdogan and Guney,2020; Kishore et al., 2018).

In fact, (Tasnim and Habiba, 2021), they conducted a study, in order to propose 7 machine learning algorithms to predict the probability of coronary heart disease, which were K-nearest neighbors' (k-NN), Random Forest (RF), Gradient Boosting, Support Vector Machine (SVM), Decision Tree (DT), Naive Bayes (NB), Logistic Regression (LR) and Neural Network (NN), for which the Hungarian dataset Cleveland and Statlog Cleveland+ from the repository of UCI machine learning, which contained 303 samples and 14 attributes, as well as Principal component analysis; among the results obtained it is shown that Random Forest achieved the best evaluation metrics, these being: Accuracy=92.85%, Recall= 91.00% y Precisión= 93.00%.

Also (Daraei and Hamidi, 2017), in their article he aimed to present a prediction model (J48) of Myocardial Infarction (MI) using data mining methods, which took into account a dataset of 455 healthy cases and 295 with myocardial infarction collected from the Shadid Madani Specialized Hospital, in Iran in 2015, as well as a hybrid method of selection of variables, which included weight by relief and genetic algorithm, where after selection the metacost classifier was applied to the dataset, where the results show that the J48 algorithm obtained an accuracy of 82.57%, sensitivity of 86.67% and F1-Score of 80.00%.

At the same time, (Mohan, Thirumalai and Srivastava, 2019), they proposed a hybrid model through the application of machine learning to improve accuracy when predicting heart disease. So within the algorithms used were: Generalized Linear model, Deep learning, Random Forest, Support Vector machine, Naive Bayes, Logistic Regression, etc., whose dataset was composed of 303 patient records, however, because there was missing data, only 297 records were worked for pre-processing, where the results show that with the combination of Random Forest (RF) and Linear Regression (RL), an accuracy of 88.7%, precisión of 90.10%, F1-Score of 90.00%, Sensitivity of 92.80% and Specificity of 82.60% are achieved.

As well as (Chowdhury et al., 2021) presented a proposed model to predict heart disease and increase the accuracy of heart disease, so they used several techniques, including: Decision Tree (DT), Logistic Regression (RL), K-Nearest Neighbors (KNN), Support Vector Machine (SVM) and Naive Bayes (NB), working with a dataset of 313 people, of which 251 were healthy, as well as the 5-Fold Cross-validation, so the results show that SVM achieved an accuracy of 91%.

Meanwhile (Kavitha et al., 2021), in their study they aimed to predict heart disease based on machine learning through an automated medical diagnosis method, for which they applied 3 ML techniques, these being: Random Forest (RF), Decision Tree (DT) and Hybrid Model, also took into consideration a dataset of 303 instances and 14 characteristics, so in the results it was evident that the hybrid model obtained the highest value of accuracy 88.70%.

Thus, in this review, he focused on the study of machine learning and deep learning methods, to carry out a deeper investigation, in order to publicize which are the most used methods as their main characteristics for the prediction of heart diseases, and in this way make technology more efficient in the health sector.

Motivation. Currently, it is common to use machine learning techniques in the field of medicine for the prediction of diseases, one of the frequent heart diseases, this because, if not treated early, it can affect the health of people, such as having complications, including: strokes, heart attacks, heart failure, etc.

Therefore, this article aims to carry out a systematic review of the research works that deal with the topic of machine learning and deep learning to predict heart disease, which allows us to know the following: In which countries have more studies been done on the use of machine learning and deep learning to predict heart disease? , What were the techniques that had the best accuracy results?, What tools are the most concurrent for the development and testing of the predictive model?, What metrics are used to determine the effectiveness of machine learning and deep learning techniques?, What type of heart disease have they been applied to? and Which selection variables are the most used in machine learning and deep learning algorithms?

This article is divided as follows: Section 2 presents the method of systematic literature review. In section 3, the respective analysis of the results according to the questions asked is presented, in section 4 the information about the discussion is presented, in section 5 shows the conclusions, along with limitations and challenges, motivations and recommendations reached in this review.

2. Research Methodology

This article is based on a systematic literature review made up of scientific databases using a PRISMA methodology (Moher et al.,2010), whose search focuses on scientific articles related to the prediction of heart disease, whose information can help save lives by providing timely information about the disease, such as existing ML and DL methods for their study.

2.1. Research Questions

The purpose of the study is to analyze, compare and synthesize studies on the prediction of heart disease using ML and LBP published from 2018 to 2023. Table 1 shows the 7 research questions (RQ) asked, with their respective motivations:

Table 1: Research questions

ID	Research Question	Motivation
Q1	In which countries have more studies been done on the use of machine learning and deep learning to predict heart disease?	Identify countries where more studies have been done on machine learning and deep learning techniques to predict heart disease
Q2	What were the most commonly used techniques?	Reveal the various machine learning and deep learning techniques being implemented to predict heart disease

Q3	What were the techniques that had the best accuracy results?	Introduce the most accurate machine learning and deep learning techniques for predicting heart disease
Q4	Which tools are the most concurrent for the development and testing of the predictive model?	Recognize machine learning and deep learning tools to predict heart disease
Q5	What metrics are used to determine the effectiveness of machine learning and deep learning techniques?	Identify the most common machine learning and deep learning metrics to predict heart disease
Q6	What kind of heart disease have they been applied to?	Identify heart disease studied with machine learning and deep learning
Q7	Which variable selection algorithms are the most used in machine learning and deep learning algorithms?	Discover the main algorithms for selecting variables related to the use of machine learning and deep learning algorithms

2.2. Search Strategy

An exhaustive search of five databases was carried out, these being: Francis and Taylor, MDPI, ScienceDirect, Scopus and Web of science, the same ones that gathered all the information and studies related to heart diseases taking into account studies published between 2018 and 2023, to carry out this search strategy the keywords were used: ("Machine learning") and ("Deep learning") and "predict" and "heart" and "disease", as shown in Figure 1 and Figure 2.

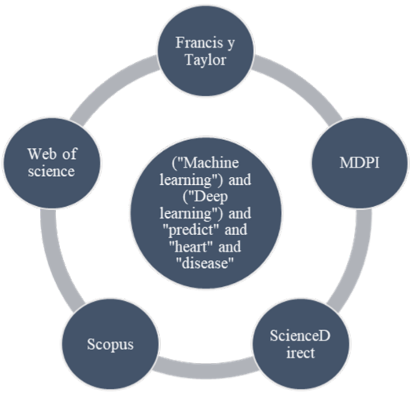

Fig. 1: Database search criteria

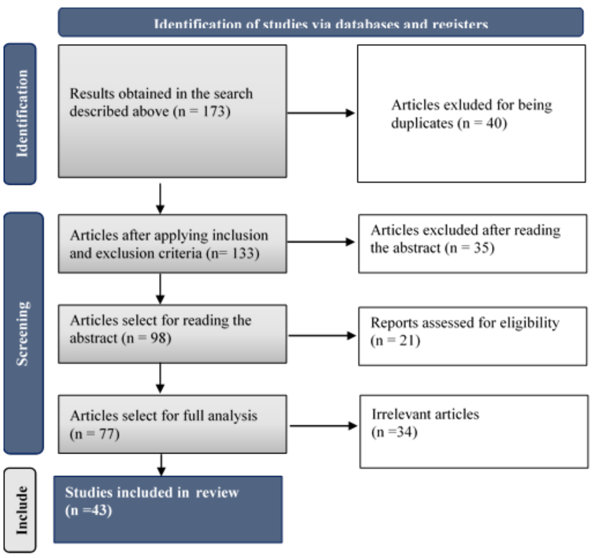

Fig. 2: PRISMA methodology for selected articles

2.3. Inclusion criteria

The following inclusion criteria was established: (i) The articles taken into account are in English language; (ii) The main focus is on the study of different techniques of Machine learning and Deep learning; (iii) All articles presented are about heart disease.

2.4. Inclusion and exclusion criteria

For the quality assessment of the study, the following exclusion criteria was defined for: (i) Articles published before 2018; (ii) Articles that study other diseases (Diabetes, cancer, etc.); (iii) Articles that are from another language.

3. Analysis of results

The results based on the 43 articles included in the systematic literature review are shown below. The following 7 sections are related to each of the research questions.

3.1. Results of countries where more studies have been done on the use of machine learning and deep learning to predict heart disease (RQ1)

Fig. 3 shows the countries that have done more studies, it can be said that there are many countries that have as a topic of interest the prediction of heart disease. The publications that have been found, being: India, Pakistan, Ethiopia, Australia, China, Egypt, Japan, South Africa, Turkey, Italy, United Kingdom, United States, Bangladesh, Saudi Arabia and Iran. However, India is the country that has the highest number of studies on diseases of this specialty (39.53%). What is evident that heart disease is a focus of interest of many researchers who within their studies consider it important to obtain information on the subject and have solutions that help identify and adequately predict cardiopathologies, this because the number of people with total CVD worldwide has doubled from 271 million in 1990 to 523 million in 2019, which means that there is an increase in deaths from 12.1 million in 1990 to 18.6 million in 2019 (Roth et al., 2020); and whose national statistics from the WHO show that CVD has a significant figure of 24% in India (World Health Organization, 2018), which is worrying, since in this country, it is evident that there are more than 10.5 million deaths per year, and cardiovascular diseases are the ones that cause 20.3% of deaths in men and 16.9% in women (Gupta, Mohan and Narula, 2016).

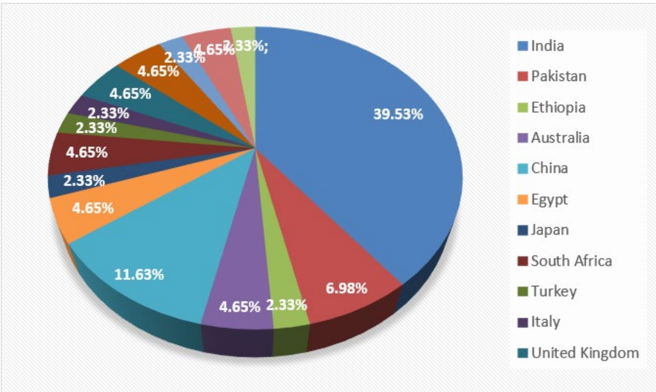

Fig. 3: Range of countries that have studied the use of ML and DL to predict heart disease

3.2. Results of the most used techniques of ML y DL (RQ2)

Within the techniques of machine learning and Deep learning most used to predict heart disease, we found 3 techniques that stand out by having the largest number of studies, these being: Random Forest, characterized by being the most used and powerful in machine learning for classification and regression problems, which collects information and builds decision trees from different samples, to finally take the average of these values, being the most efficient to predict heart disease (Kompella et al.,2022), so it is evident that it was the most used by most authors, that is, by 9.18%; at the same time as SVM, it is a machine learning algorithm for categorization, whose purpose is to locate the optimal separation hyperplane for the unobserved sequences with the lowest possible classification error (Kumar and Priyadarsini, 2022; Sandhya,2020), by which 7.65% of the authors use it together with Logistic Regression, with 15 articles these last 2, as shown in Fig. 4. This shows that the techniques detailed above are the most efficient for predicting cardiac pathologies.

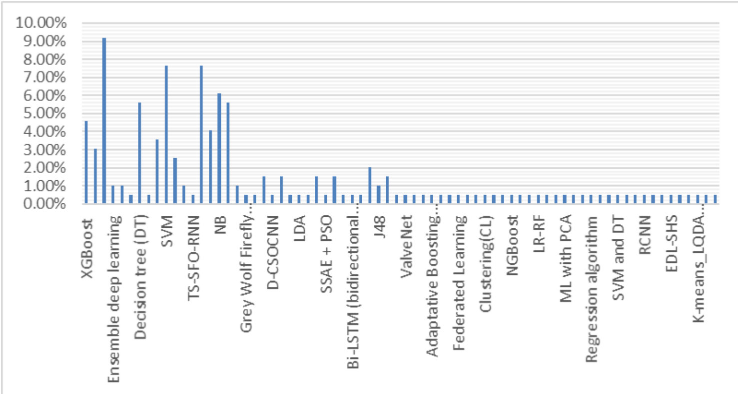

Fig. 4: Machine learning and deep learning techniques to predict heart disease

3.3. Results of Machine learning and Deep learning techniques with best accuracy(RQ3)

The results analyzed with respect to the techniques that had better accuracy scores in predicting heart disease were: Random Forest, Deep learning, SVM, XGBoost, where 16.28% of the authors make use of these algorithms, also 4.25% of the studies showed Staking Ensemble as the best algorithm, while 9.30% of the authors showed that KNN and NN obtained better precisions, whose accuracy ranges from 90%-100%, as shown in Table 2. It should be noted that this is related to the size of the data taken into account in the investigations (Gupta et al., 2022).

Table 2: Precision of Machine Learning and Deep Learning Techniques to Predict Heart Disease

Ref.	Technique	Precision
Simegn, Gebeyehu and Degu (2022)	XGBoost	98.00%
	ANN	95.00%
	Random forest	100.00%
Sheeba et al (2022)	Ensemble deep learning	----
	Support vector machine (SVM)-Radi	----
	SVM-DT	----
	Decision tree (DT)	----
	Mixed kernel extreme learning machine (MKELM)	88.20%
Wankhede, Sambandam and Kumar (2022)	Ensemble deep learning	98.33%
Ganesh and Nithiyanantham (2022)	NN	51.51%
	SVM	54.05%
	CNN	59.25%
	RNN	54.05%
	TS-SFO-RNN	69.23%
Ayon, Islam and Hossain (2022)	SVM	95.00%
	LR	91.00%
	DNN	92.00%
	DT	91.00%

	NB	91.00%
	Random Forest	90.00%
	K-NN	93.00%
Malathi et al (2021)	AlexNet	94.00%
Jiang et al (2021)	LR	87.20%
	SVM	86.80%
	KNN	84.50%
	RF	84.00%
	NB	79.10%
	DT	77.00%
	XGB	80.40%
Deepika and Balaji (2022)	Grey Wolf Firefly algorithm with Differential Evolution (GF-DE)	99.29%
Hussain et al (2022)	KNN	80.00%
	SVM-L	97.14%
	NB	88.57%
	NN	88.57%
	GBM	91.42%
	XGB	88.57%
Chamundeshwari, Biradar and Udaykumar (2022)	D-CSOCNN	97.45%
Almulihi et al (2022)	Staking ensemble	97.42%
	LR	83.00%
	CNN-LSTM	86.10%
Kanda et al (2022)	Extreme gradient boosting	14.10%
Mienye and Sun (2021)	KNN	60.80%
	LR	79.00%
	LDA	80.40%
	SVM	80.00%
	Decision tree	69.99%
	Softmax classifier	71.50%
	XGBoost	86.40%
	Random forest	91.40%
	AdaBoost	93.20%
	Proposed SSAE + PSO	93.00%
Nancy et al (2022)	LSTM	95.07%
	FLSTM	98.03%
	Bi-LSTM (bidirectional long short-term memory) (Proposed)	98.90%
Umer et al (2022)	VGG16	90.00%
	AlexNet	90.00%
	CNN	94.00%

Ozcan and Peker (2023)	CART model	88%
Brunese et al (2020)	Deep learning network	98.99%
	J48	71.00%
	MLP	92.00%
	UCL	58.00%
Chang, Bhavani and Xu (2022)	Random Forest (RF)	83.00%
Elias et al (2022)	ValveNet	32.00%
	Gradient Boosting (GBM)	86.80%
	Random Forest (RF)	87.80%
	Multilayer perceptron (MLP)	88.30%
	k-Nearest Neighbor(KNN)	90.40%
	Extra Trees Classifier(ET)	90.20%
	Extreme Gradient Boosting(XGB)	85.50%
Mohapatra et al (2023)	Support Vector Classifier (SVC)	88.80%
	Support Vector Classifier (SGD)	94.70%
	Adaptive Boosting(ADB)	76.00%
	Decision Tree (Cart)	81.00%
	Stacked	92.60%
Mpanya et al (2021)	Random forest	81.70%
	Naïve Bayes	85.70%
	Decision trees	73.50%
	Gradient boosting	73.80%
	SVM	69.00%
	Logistic regression	62.60%
	CART	63.60%
Roy, Roy and Mandal (2022)	KNN	75.00%
	SVM	98.00%
	RF	98.00%
	NB	95.00%
	ANN	56.00%
Shrivastava, Sharma and Kumar (2023)	CNN–Bi-LSTM (Propose Method)	96.84%
	LR (Logistic Regression)	90.13%
	SVM (Support Vector Machine)	88.35%
	RF (Random Forest)	85.85%
	DT (Decision Tree)	80.37%
	HRFLM	86.77%
Uddin and Halder (2021)	Proposed MLDS (Multilayer dynamic system)	97.00%
	XGB	75.72%
	LR	72.26%
	SVM	77.55%
	KNN	69.62%
	DT	63.10%

Yan et al (2022)	XGBoost	85.70%
	LightGBM	85.30%
	Random Forest	81.20%
	NGBoost	73.50%
	Logistic regression	60.30%
	MLP	58.30%
Sajja and Kalluri (2020)	Logistic Regression	86.83%
	Naive Bayes	77.04%
	KNN	68.86%
	SVM (Linear)	85.29%
	SVM (RBF)	81.96%
	Neural Network	86.97%
	CNN (Proposed Network)	94.78%
Waqar et al (2021)	SMOTE-based AdaBoost	81.30%
	SMOTE-based bagging	81.70%
	SMOTE-based random forest	85.10%
	SMOTE-based KNN	82.50%
	SMOTE-based logistic regression	81.80%
	SMOTE based na¨ıve Bayes	81.90%
	SMOTE-based support vector machine	84.50%
	SMOTE-based vote	82.30%
	SMOTE-based artificial neural network	96.10%
Alotaibi and Alzahrani (2022)	NB	84.00%
	LR-RF	75.00%
	AdaBoost	77.00%
	LR	73.00%
	RF	76.00%
	SVM	79.00%
Alqahtani et al (2022)	RF	90.03%
	KNN	87.53%
	DT	86.23%
	XGB	88.25%
	DNN	97.77%
	KDNN	98.52%
	ML ensemble	88.02%
	Stacked classifier	87.32%
Swain et al (2022)	Logistic Regression	95.35%
	Support Vector Machine (SVM)	91.86%
	MLP with PCA	91.86%
	Deep neural network	93.02%
	Bootstrap aggregation with random forests	97.67%
Albert, Murugan and Sripriya (2022)	BOML algorithm	75.00%
Ordikhani et al (2022)	Regression algorithm	73.90%
	Naive Bayes Classifier	72.70%

	Random Forest	71.20%
	Network Configuration	73.70%
	Deep Learning	74.00%
Huang et al (2022)	DT	95.00%
	SVM and DT	86.72%
	RF	83.26%
	KNN	97.77%
	NN	97.70%
	Adaptive multi-layer networks (AMLN)	97.80%
Damodharan and Goel (2022)	DBN	90.03%
	RCNN	91.70%
	AWMYolov4	95.60%
Bangare et al (2022)	CART model	93.30%
Elwahsh et al (2021)	SVM	95.00%
	Logistic regression	90.00%
	Decision tree	91.00%
	Naïve Bayes	89.80%
	ANN	93.00%
	SHDML	98.00%
Guo et al (2021)	LSTM	62.00%
	DNN	59.00%
	RF	60.00%
	LR	58.00%
	NB	60.00%
Barbieri et al (2022)	Deep Learning network(DNN)	46.80%
Pan et al (2020)	ANN	84.30%
	DNN	88.50%
	EDL-SHS	89.20%
	RNN	90.20%
	NN	93.10%
	EDCNN	97.20%
Saikumar et al (2022)	SVM	55.00%
	CNN	59.00%
	FGCNet	63.00%
	K-means_LQDA _DG_ConvoNet	76.00%
You et al (2023)	Expert feature	14.50%
	CNN	22.96%
	LSTM	19.63%
	LR	13.37%
	Deep neural network	21.06%
Wang et al (2022)	Neural network	74.00%
	InceptionV3	64.14%
	Random Forest	75.79%

3.4. Results of Machine learning and Deep Learning tools (RQ4)

Of the 43 articles analyzed, it is observed that 27 studies use Python (49.09%) and 7 Keras studies (12.73%), which shows that these are the most recurrent tools for the development and testing of the predictive model in the prediction of heart disease as detailed in Table 3. Both have very similar advantages such as: the number of free and open source libraries, possibility of extracting data from repositories, simple language and great potential (Kazil, Masad and Crooks, 2020; Chicho and Sallow, 2021).

Table 3: ML and DL tools to predict heart disease

Ref.	Tool	Related studies	Quantity (%)
H01	Matlab	(Sheeba et al., 2022),(Ganesh, 2022), (Swathy and Saruladha, 2022)	3 (5.45%)
H02	Keras	(Almulihi et al., 2022), (Umer et al.,2022), (Brunese et al., 2020), (Elias et al., 2022), (Roy, Roy and Mandal, 2022), (Waqar et al., 2021), (Saikumar et al., 2022)	7 (12.73%)
H03	Pytorch	(Elwash et al.,2021), (Barbieri et al., 2022)	2 (3.64%)
H04	Python	(Wankhede, Sambandam and Kumar, 2022), (Ayon, Islam and Hossain, 2022), (Malathi et al,2021), (Jiang et al.,2021), (Deepika and Balaji, 2022),(Almulihi et al., 2022), (Kanda et al., 2022), (Mienye and Sun, 2021), (Brunese et al., 2020), (Chang et al., 2022), (Elias et al.,2022), (Mohapatra et al.,2023), (Roy, Roy and Mandal, 2022), (Shrivastava, Sharma and Kumar ,2023), (Swathy and Saruladha, 2022), (Uddin and Halder, 2021), (Yan et al.,2022), (Waqar et al., 2021), (Alqahtani et al., 2022), (Albert, Murugan and Sripriya, 2022), (Ordikhani et al., 2022), (Damodharan and Goel, 2022), (Elwahsh et al., 2021), (Guo et al., 2021),(Barbieri et al.,2022),(Saikumar et al., 2022),(You et al.,2023)	27 (49.09%)
H05	TensorFlow	(Almulihi et al., 2022), (Nancy et al., 2022), (Umer et al., 2022), Brunese et al., 2020), (Waqar et al., 2021)	5 (9.09%)
H06	Weka	(Ayon, Islam and Hossain, 2022), (Swathy and Saruladha, 2022), (Uddin and Halder, 2021), (Alotaibi and Alzahrani, 2022)	4 (7.27%)
H07	R	(Malathi et al., 2021), (Ozcan and Peker, 2023)	2 (3.64%)
H08	Tangara	(Swathy and Saruladha, 2022)	1 (1.82%)
H09	Anaconda	(Waqar et al., 2021), (Albert, Murugan and Sripriya, 2022), (Elwash et al., 2021)	3 (5.45%)
H10	Keel	(Alotaibi and Alzahrani, 2022)	1 (1.82%)

3.5. Results of the most common Machine learning and Deep learning metrics (RQ5)

Fig.5 shows the different metrics used to predict heart disease, in which is the Accuracy, which represents the total percentage of all predictions that are correctly classified [81], where it is shown that this metric was the most used by most authors (15. 35%) ., at the same time of Accuracy, represented by the number of successful instances recovered divided by the totality of recovered instances, classified as positive predictive value (PPV) and false positive value (FPV) used in medical systems [82], in which 13.60% of the authors use it to determine the effectiveness of the algorithms, finally the F1-Score represented as the combination of the Recall measures and accuracy in a single value [83], evidenced by 10.96% of authors who use it. This shows that these metrics are the most reliable and intuitive, as it is also important to calculate the relationship between the correct predictions of a model versus the total predictions that were carried out, in order to review how the algorithm performs.

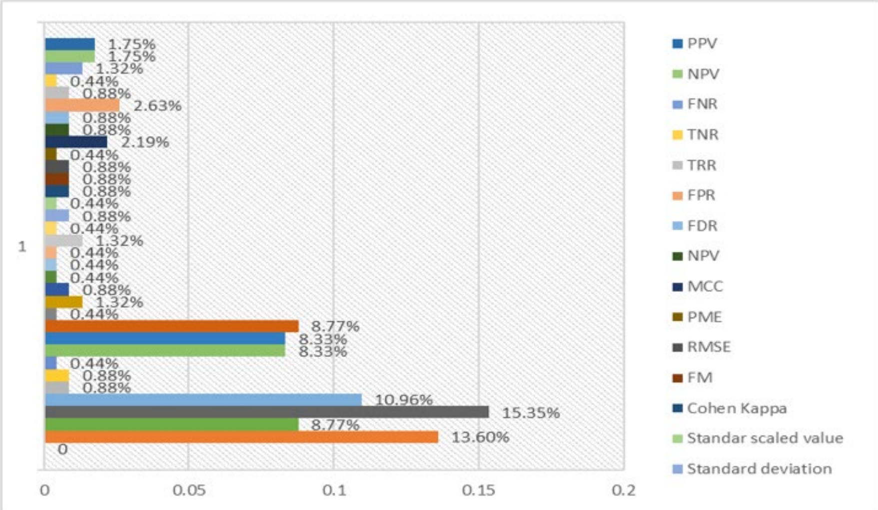

Fig. 5: Machine Learning and Deep Learning Metrics to Predict Heart Disease

3.6. Results of kind of heart disease have been applied (RQ6)

Fig.6 shows Coronary Artery Heart is the type of heart disease that has been applied most frequently, this is because most authors (20.41%) and Ischaemic heart disease (7.14%), see the need to study this pathology (CAH), even more so when this is the most outstanding and remains one of the main reasons for deaths, while according to the WHO report indicates that 17.3 million deaths from heart disease, 7.3 million are coronary heart disease [2], and since the disease is a difficult procedure to predict, it is important that patients require early and adequate diagnoses of CAH for effective treatments; in addition to the approach to the use of technology [73, 74]. However, there are different heart disease that have been investigated by 1 author representing 30.61%.

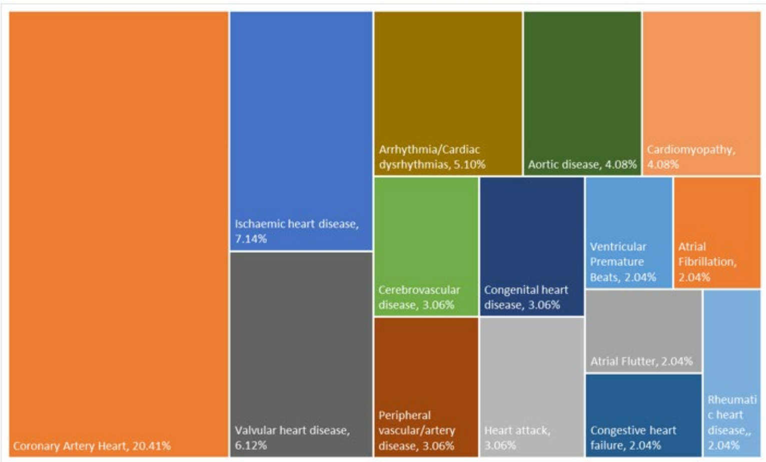

Fig. 6: Type of heart disease most common applied with Machine Learning and Deep Learning

3.7. Results of the Variable selection algorithms used in machine learning and deep learning algorithms (RQ7)

To discover the main variable selection algorithms, Kernel, Kernel, interpreted as a function that returns the inner product between the images of the inputs based on their characteristics, where its purpose is to detect nonlinear relationships between algorithms [84], so this is the most used by most authors (31.71%), however, other authors chose to use different ones, such as Information Gain, represented as an entropy-based feature Evaluation method, used in machine learning to measure the amount of the data that a function provides about the target class [85], so 9.76% of the authors make use of this variable selection algorithm. Making use of these variable selection algorithms helps the techniques used within the studies to have better results with respect to the prediction of heart disease and to be more accurate, as shown in Table 4.

Table 4: Variable selection algorithms using Machine Learning and Deep Learning

Ref	Variable selection	Related studies	Quantity(%)
SV01	Information Gain	(Deepika and Balaji, 2022),(Mpanya et al., 2021),(Uddin and Halder, 2021), (Ordikhani et al.,2022)	4 (9.76%)
SV02	Kernel	(Simegn, Gebeyehu and Degu, 2022),(Sheeba et al.,2022),(Ayon, Islam and Hossain,2022),(Malathi et al.,2021),(Jiang et al.,2021),(Elias et al.,2022),(Mohapatra et al.,2023),(Roy, Roy and Mandal,2022),(Waqar et al.,2021), (Swain et al.,2022),(Damodharan and Goel,2022),(Pan et al.,2020),(You et al.,2020)	13 (31.71%)
SV03	GainRatio-AttributeEval	(Uddin and Halder,2021)	1 (2.44%)
SV04	ReliefFAttributeEval	(Deepika and Balaji,2022),(Nancy et al.,2022),(Alotaibi and Alzahrani,2022)	3 (7.32%)
SV05	Channel selection algorithm (CSA)	(Ganesh and Nithiyanantham,2022)	1 (2.44%)
SV06	Motivated Feature Selection (MFS)	(Deepika and Balaji,2022)	1 (2.44%)
SV07	mrMR	(Deepika and Balaji,2022)	1 (2.44%)
SV08	Recursive Feature Elimination (RFE)	(Almulihi et al.,2022),(Sajja and Kalluri,2020)	2 (4.88%)
SV09	Sequential forward selection	(Nancy et al.,2022)	1 (2.44%)
SV10	chroma stft	(Brunese et al.,2020)	1 (2.44%)
SV11	spectral centroid	(Brunese et al.,2020)	1 (2.44%)
SV12	spectral bandwidth	(Brunese et al.,2020)	1 (2.44%)
SV13	zero crossing rate	(Brunese et al.,2020),(Roy, Roy and Mandal,2022)	2 (4.88%)
SV14	Mel-Frequency Cepstral Coefficients(MFCC)	(Brunese et al.,2020)	1 (2.44%)
SV15	MCC feature subset selection	(Mohapatra et al.,2023)	1 (2.44%)
SV16	Sequential Backward Selection (SB S)	(Mohapatra et al.,2023)	1 (2.44%)
SV17	Correlation Attribute Evaluator (CAE)	(Uddin and Halder,2021)	1 (2.44%)
SV18	Chi-squared feature selection	(Alotaibi and Alzahrani,2022)	1 (2.44%)
SV19	Forward Feature Selection	(Alotaibi and Alzahrani,2022), (Ordikhani et al.,2022)	2 (4.88%)
SV20	ROI Selection	(Damodharan and Goel,2022)	1 (2.44%)

SV21	Fast Correlation - Based feature Selection(FCBF)	(Damodharan and Goel,2022)	1 (2.44%)

4. Discussion

4.1. General Findings

In the present literature review it is evident that there are a variety of articles regarding the prediction of heart disease using Machine learning (ML) and Deep learning (DL). The initial consultation based on the required information obtained a quantity of 173 results, however, of these exclusion criteria had to be made, since many were duplicates and did not comply with the request, which finally remained in 43 articles. From the theoretical perspective, ML and DL techniques can provide advantages to efficiently analyze and predict heart pathologies, this due to the tools and information provided.

The research question 1 of this study is with respect to the countries where more articles have been made about Machine learning and Deep learning techniques to predict heart disease, so Fig.3 shows the range of countries, mainly the countries of Asia were the ones that obtained the highest number of studies investigated. The country that conducted the most research was India, followed by China and Pakistan. The country of India shows a high percentage of studies, where the number of people who die in that place due to cardiovascular diseases have increased and it is necessary and important to obtain information and solutions that help reduce that figure, however China and Pakistan have a very low amount compared to India, which shows that although studies are done, they still need to be strengthened and a more detailed analysis carried out.

With question 2, we wanted to reveal the various techniques of machine learning and deep learning that are implemented to predict heart disease. The number of articles found within the review showed a variety of techniques, whose results show that 18 articles consider Random Forest as the most used, followed by 15 articles about SVM. In contrast to this, 2 articles use Logistic Regression to predict heart disease, demonstrating the efficiency of such techniques to detect pathologies.

At the same time, referring to question 3, when talking about the techniques that obtained better precision scores, the results are very confusing with respect to the techniques treated, so the techniques are detailed in a general way and of them several are repeated. As, for example, in 7 articles they deal with four different techniques (Random Forest, Deep learning, SVM and XGBoost), following with KNN and NN with 4 articles, so it should be noted that (Brunese et al., 2020) is present with several techniques Deep learning, KNN, Stacking and NN. This could be due to the fact that said author conducts a detailed study with the intention of providing a more complete analysis, based on the information required and the size of the same.

With respect to research question 4, when it comes to the most concurrent machine learning and deep learning tools there is a significant difference regarding the number of studies. In Table 3, for example, it can be seen that 16 articles detail that Python is the most used tool, on the other hand, it is shown that 7 studies apply Keras. So both tools are very useful for the advantages they have, which highlights the number of free libraries, their potential and easy to code language.

In research question 5, the most common metrics that were used were identified, whose importance lies in the fact that these are reliable and intuitive and help to have the most accurate prediction and improve the performance of the algorithm. This was seen in 31 articles, where Accuracy was the one that obtained the most research, similar case with the metrics of Precision and F1-Score, which shows that the authors take into account these precise and necessary indicators.

Also, in research question 6, we sought to identify the heart diseases studied with ML and LBP, whose most applied disease was Coronary Artery Heart with 31 articles, the large number of studies is due to the fact that this pathology is the most worrying due to the number of deaths (7.3 million) and it is very difficult to predict. However, in contrast to what has been said, only 7 articles talk about another type of disease: Ischaemic heart disease, which based on the low number of articles found within the study, should emphasize the study of this cardiac pathology, since it is also important to be caused by the narrowing of the arteries of the heart, that is, blood vessels and is harmful to the health of patients.

Finally, the main variable selection algorithms, in Table 4, on the one hand, 13 studies show that Kernel is the most used, but some authors also take into account both Information Gain and ReliefFAttributeEval, being conceivable that by making use of these better precision results are obtained.

5. Conclusions

The systematic review shows the current context of the study regarding the use of Machine Learning and Deep Learning for the prediction of heart disease. So the grouping of 43 articles analyzed indicates that this is a topic of interest within the health sector, which is in the process of expansion in the coming years, however, today the total number of studies related to different specific pathologies of the heart is not expected, since many of these still need to be treated to carry out a more detailed study. Identifying the countries that have done more studies of the use of Machine learning and Deep learning is necessary, since in this way it is known which are the places that researchers pay greater interest (India, China and Pakistan) as their importance to obtain information, and in this way reduce deaths when making use of technology. Whose most used techniques (Random Forest, SVM and Logistic Regression) are very promising, because they prove to be the most efficient when predicting diseases. At the same time, XGBoost, Ensemble Deep learning, etc., were the ones that obtained a better score, between 90%-100%. The use of Machine learning and Deep learning tools has advantages, being the case of Python which is open source, its simple language and information extraction. So the most common metrics (Accuracy, Accuracy and F1-Score) together with the main variable selection algorithms (Kernel, Information Gain and ReliefFAttributeEval) were important to optimize the performance of the algorithm and predict heart disease, including the most applied in this branch: Coronary Artery Heart. Therefore, the use of Machine Learning and Deep Learning techniques helps predict heart disease.

5.1. Research challenges, limitations and Future research directions.

Within the study, only 5 databases were taken into account, these being the most recognized: Francis and Taylor, MDPI, ScienceDirect, Scopus and Web of Science, therefore, there may be many articles indexed in other sources that were not located for research and that could have been related to the topic. Likewise, publication bias plays an important role in this type of studies (literature reviews).

Although it was possible to answer the research questions as posed, the research has some limitations, first of all, most of the studies deal with 1 type of heart disease: Coronary Artery Heart, however, not many were found that talk about the other pathologies treated in the study, such as: Ischaemic heart disease, Valvular heart disease, Arrhythmia / Cardiac dysrhythmias, etc., which could be summarized in a high percentage of unsuccessful articles in the research area. Likewise, it was observed that when making use of Deep Learning in the health sector, there are still challenges regarding the volume, quality, availability and standardization of data (Chamundeshwari, Biradar and Udaykumar, 2022), while some algorithms such as decision tree face challenges such as parameter adjustment, biased learned trees and overfting (Albert, Murugan and Sripriya, 2022).

On the other hand, the models must be updated on a regular basis so that it is available in decision making, ensuring the privacy and security of information, being a significant challenge and high cost, in addition to interpreting the data for doctors is not easy (Huang et al., 2022).

5.2. Recommendations

Based on the large amount of information related to heart disease worldwide, it is recommended to integrate all data and apply machine learning techniques with big data. Likewise, the construction of graphical interfaces with machine learning that allow medical personnel to help better decision making is recommended.

Acknowledgements

Please acknowledge collaborators or anyone who has helped with the paper at the end of the text.

References

Ahsan M.M., Mahmud M.P., Saha P.K., Gupta K.D. & Siddique, Z. (2021). Effect of data scaling methods on machine learning algorithms and model performance. *Technologies,* 9(3), 52.

Albert A. J., Murugan R. & Sripriya T. (2022). Diagnosis of heart disease using oversampling methods and decision tree classifier in cardiology. *Research on Biomedical Engineering*, 1-15.

Aljefree N.M., Shatwan I.M. & Almoraie, N.M. (2021). Association between nutrients intake and coronary heart disease among adults in saudi arabia: A case-control study. *Progress in nutrition*, 23(3),1-12.

Almeida G. & Tavares J. M. (2020). Deep Learning in Radiation Oncology Treatment Planning for Prostate Cancer: A Systematic Review. *Journal of Medical Systems*, 44(10),179.

Almulihi A., Saleh H., Hussien A.M., Mostafa S., El-Sappagh S., Alnowaiser K., Ali A.A. & Refaat H. M. (2022). Ensemble Learning Based on Hybrid Deep Learning Model for Heart Disease Early Prediction. *Diagnostics*, 12 12), 3215.

Alonso S. G., De La Torre I., Hamrioui S., López M., Barreno D. C., Nozaleda L. M. & Franco M. (2018). Data mining algorithms and techniques in mental health: a systematic review. *Journal of Medical Systems*, 42(9), 1-15.

Alotaibi N. & Alzahrani M. (2022). Comparative analysis of machine learning algorithms and data mining techniques for predicting the existence of heart disease. *International Journal of Advanced Computer Science and Applications*, 13(7), 1-9.

Alqahtani A., Alsubai S., Sha M., Vilcekova L. & Javed T. (2022). Cardiovascular Disease Detection using Ensemble Learning. *Computational Intelligence and Neuroscience*, 1-9.

Angraal S., Mortazavi B.J., Gupta A., Khera R., Ahmad T., Desai N.R., Jacoby D.L., Mausidi F.A., Spertus J.A. & Krumholz H.M. (2019). Machine learning prediction of mortality and hospitalization in heart failure with preserved ejection fraction. *JACC: Heart Failure*, 8(1), 12–21.

Ayon S. I., Islam M. & Hossain R. (2022). Coronary Artery Heart Disease Prediction: A Comparative Study of Computational Intelligence Techniques. *IETE Journal of Research*, 68(4), 2488-2507.

Bangare S. L., Virmani D., Karetla, G. R., Chaudhary P., Kaur H., Bukhari A.N. & Miah S. (2022). Forecasting the Applied Deep Learning Tools in Enhancing Food Quality for Heart Related Diseases Effectively: A Study Using Structural Equation Model Analysis. *Journal of Food Quality*,1-8.

Barbieri S., Mehta S., Wu B., Bharat C., Poppe K., Jorm L. & Jackson R. (2022). Predicting cardiovascular risk from national administrative databases using a combined survival analysis and deep learning approach. *International Journal of Epidemiology*, 51(3), 931.

Berno M., Canil M., Chiarello N., Piazzon L., Berti F., Ferrari F., Zaupa A., Ferro N., Rossi M. & Susto G.A. "A machine learning-based approach for advanced monitoring of automated equipment for the entertainment industry," *IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4. 0&IoT)*, 2021.

Bhowmick A., Mahato K.D., Azad C. & Kumar U. "Heart Disease Prediction Using Different Machine Learning Algorithms," *IEEE World Conference on Applied Intelligence and Computing (AIC)*, 2022.

Brunese L., Martinelli F., Mercaldo F. & Santone A. (2020). Deep learning for heart disease detection through cardiac sounds. *Procedia Computer Science*, 176, 2202-2211.

Chamundeshwari, Biradar N., Udaykumar. (2022). Adaptive Despeckling and Heart Disease Diagnosis by Echocardiogram using Optimized Deep Learning Model. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 1-17.

Chang V., Bhavani V. R., Xu A. Q. & Hossain M. A. (2022). An artificial intelligence model for heart disease detection using machine learning algorithms. *Healthcare Analytics*, 2, 100016.

Chicco D. & Jurman G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC genomics*, 21, 1-13.

Chicho B. T. & Sallow A. B. (2021). A Comprehensive Survey of Deep Learning Models Based on Keras Framework. *Journal of Soft Computing and Data Mining*, 2(2), 49-62.

Chowdhury M. N., Ahmed E., Siddik M. A. & Zaman A. U. "Heart disease prognosis using machine learning classification techniques," *International Conference for Convergence in Technology (I2CT)*, 2021.

Dalianis H. & Dalianis H. 2018. Evaluation metrics and evaluation. Clinical text mining: secondary use of electronic patient records. Springer, NY, 45-53.

Damodharan D. & Goel, A. K. (2022). An Implementation of Cardiovascular Disease Prediction in Ultrasonography Images Using AWMYOLOv4 Deep Learning Mode. *International Journal on Recent and Innovation Trends in Computing and Communication*, 10(9),40-52.

Daraei A. & Hamidi H. (2017). An efficient predictive model for myocardial infarction using cost-sensitive j48 model. *Iran Journal of Public Health*, 46(5), 682-692.

Daza A., Guerra C., Cervera N. & Burgos, E. (2022). Predicting Academic Performance through Data Mining: A Systematic Literature. *TEM Journal*, 11(2), 939-949.

Daza A., Guerra C., Cervera N. & Burgos, E. (2022). Predicting Academic Performance using a Multiclassification Model: Case Study. *International Journal of Advanced Computer Science and Applications*, 13(9), 1-9.

Deepika D. & Balaji, N. (2022). Effective heart disease prediction with Greywolf with Firefly algorithm-differential evolution (GF-DE) for feature selection and weighted ANN classification. *Computer Methods in Biomechanics and Biomedical Engineering*, 25(12), 1409-1427.

Elias P., Poterucha T. J., Rajaram V., Moller L.M., Rodriguez V., Bhave s. Hahn R.T., Tison G., Abreau S.A., Barrios J., Torres J.N., Hughes J.W., Perez M.V., Finer J., Kodali S., Khalique O., Hamid N., Schwartz A., Homma S. Kumaraiah D. & Perotte A.J. (2022). Deep learning electrocardiographic analysis for detection of left-sided valvular heart disease. *Journal of the American College of Cardiology*, 80(6), 613-626.

Elwahsh H., El-Shafeiy E., Alanazi S. & Tawfeek M. A. (2021). A new smart healthcare framework for real-time heart disease detection based on deep and machine learning. *PeerJ Computer Science*, 7, 1-34.

Erdogan A. & Guney S. “Heart Disease Prediction by Using Machine Learning Algorithms,” *Signal Processing and Communications Applications Conference (SIU)*, 2020.

Ferreras A., Sumalla S., Martínez R., Elío I., Tutusaus K., Prola T., Vidal J.L., Sahelices B. & Torre I. (2023). Systematic Review of Machine Learning applied to the Prediction of Obesity and Overweight. *Journal of Medical Systems*, 47(8),1-11.

Ganesh V. M. & Nithiyanantham J. (2022). Heuristic-based channel selection with enhanced deep learning for heart disease prediction under WBAN. *Computer Methods in Biomechanics and Biomedical Engineering*, 25(13), 1429-1448.

Guo A., Smith S., Khan Y. M., Langabeer J. R. & Foraker R. E. (2021). Application of a time-series deep learning model to predict cardiac dysrhythmias in electronic health records. *Plos one*, 16(9), 1-13.

Gupta R., Mohan I. & Narula J. (2016). Trends in coronary heart disease epidemiology in India. *Annals of Global Health*, 82(2), 307–315.

Gupta S., Saluja K., Goyal A., Vajpayee A. & Tiwari V. (2022). Comparing the performance of machine learning algorithms using estimated accuracy. *Measurement: Sensors*, 24, 100432.

Hall J.E., Hall M.E. 2016. Guyton and Hall Textbook of Medical Physiology e-Book. Elsevier, Amsterdam, 1152.

Huang J. D., Wang J., Ramsey E., Leavey G., Chico T.J. & Condell J. (2022). Applying Artificial Intelligence to Wearable Sensor Data to Diagnose and Predict Cardiovascular Disease: A Review. *Sensors*, 22, 8002.

Hussain L., Lone K. J., Awan I. A., Abbasi A. A. & Pirzada J. R. (2022). Detecting congestive heart failure by extracting multimodal features with synthetic minority oversampling technique (SMOTE) for imbalanced data using robust machine learning techniques. *Waves in Random and Complex Media*, 32(3), 1079-1102.

JayaSree M. & Rao LK. (2020). WITHDRAWN: Survey on - identification of coronary artery disease using deep learning. *Materials Today: Proceedings*,1-1.

Jiang Y., Zhang, X., Ma, R., Wang X., Liu J., Keerman M., Yan Y., Ma J., Song Y., Zhang J., He J., Guo S. & Guo H. (2021). Cardiovascular Disease Prediction by Machine Learning Algorithms Based on Cytokines in Kazakhs of China. *Clinical Epidemiology*, 417-428.

Justesen N., Bontrager P., Togelius J. & Risi, S. (2019). Deep learning for video game playing. *IEEE Transactions on Games*, 12 (1), 1-20.

Kanda E., Okami S., Kohsaka S., Okada M., Ma X., Kimura T., Shirakawa K. & Yajima T. (2022). Machine Learning Models Predicting Cardiovascular and Renal Outcomes and Mortality in Patients with Hyperkalemia. *Nutrients*, 14(21), 4614.

Kavitha M., Gnaneswar G., Dinesh R., Sai Y. R. & Suraj R. S. “Heart disease prediction using hybrid machine learning model,” *International conference on inventive computation technologies (ICICT)*, 2021.

Kazil J., Masad D. & Crooks A. “Utilizing python for agent-based modeling: The mesa framework,” *Social, Cultural, and Behavioral Modeling: 13th International Conference*, 2020.

Kishore A., Kumar A., Singh K., Punia M. & Hambir Y. (2018). Heart attack prediction using deep learning. *International Research Journal of Engineering and Technology (IRJET)*, 5(4), 2395-0072.

Kompella S. C. & Kolluru S.S. (2022). Heart Disease Prediction Using Random Forest Algorithm. *International Research Journal of Engineering and Technology (IRJET)*, 9(3), 1191-1194.

Kumar B. M. & Priyadarsini P. S. (2022). Efficient Prediction of Heart Disease using SVM Classification Algorithm and Compare its Performance with Linear Regression in Terms of Accuracy. *Journal of Pharmaceutical Negative Results*, 13(4), 1430-1437.

Lip Y.H., Collet J.P., De Caterina R., Fauchier L., Lane D.A., Larsen T.B., Marin F.,Morais J., Narasimhan C., Olshansky B., Pierard L., Potpara T., Sarrafzadegan N., Sliwa K., Varela G., Vilahur G., Weiss T., Boriani G. & Rocca B. (2017). Antithrombotic therapy in atrial fibrillation associated with valvular heart disease: a joint consensus document from the european heart rhythm association (ehra) and european society of cardiology working group on thrombosis, endorsed by the esc working group on valvular heart disease, cardiac arrhythmia society of southern africa (cassa), heart rhythm society (hrs), asia pacific heart rhythm society (aphrs), south african heart (sa heart) association and sociedad latinoamericana de estimulacion cardiaca y electrofisiologia (soleace). *Ep Europace*, 117(12),2215-2236.

Malathi S., Arockia R.Y, Abhishek K., Kumar A., Kumar A., Elangovan D., Kumar A., Chitra B & Abirami A. (2021). Prediction of cardiovascular disease using deep learning algorithms to prevent COVID 19. *Journal of Experimental & Theoretical Artificial Intelligence*, 1-16.

Mendo I.R., Marques G., De la Torre Díez, I., López M. & Martín F. (2021). Machine Learning in Medical Emergencies: a Systematic Review and Analysis. *Journal of Medical Systems*, 45(10), 88.

Mienye I.D. & Sun Y. (2021). Improved Heart Disease Prediction Using Particle Swarm Optimization Based Stacked Sparse Autoencoder. *Electronics*,10(19), 2347.

Mohan S., Thirumalai C. & Srivastava G. (2019). Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques. *IEEE Access*,7, 81542–81554.

Mohapatra S., Maneesha S., Mohanty S., Patra P.K., Bhoi S.K., Sahoo K.S. & Gandomi A.H. (2023). A stacking classifiers model for detecting heart irregularities and predicting Cardiovascular Disease. *Healthcare Analytics*, 3, 100133.

Moher D., Liberati A., Tetzlaff J. & Altman D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, 8(5), 336–341.

Morris S. A. & Lopez K. N. (2021). Deep learning for detecting congenital heart disease in the fetus. *Nature Medicine*, 27, 764–765.

Mpanya D., Celik T., Klug E. & Ntsinjana H. (2021). Predicting mortality and hospitalization in heart failure using machine learning: A systematic literature review. *IJC Heart & Vasculature*, 34, 100773.

Nancy A.A., Ravindran D., Raj P.M., Srinivasan K. & Gutierrez D. (2022). IoT-Cloud-Based Smart Healthcare Monitoring System for Heart Disease Prediction via Deep Learning. *Electronics*, 11(15), 2292.

Ordikhani M., Saniee A. M., Prugger C., Prugger C., Hassannejad R., Mohammadifard N. & Sarrafzadegan N. (2022). An evolutionary machine learning algorithm for cardiovascular disease risk prediction. *PLoS ONE*, 17(7), 1-16.

Ozcan M. & Peker S. (2023). A classification and regression tree algorithm for heart disease modeling and prediction. *Healthcare Analytics*, 3, 100130.

Pan Y., Fu M., Cheng B., Tao X. & Guo J. (2020). Enhanced deep learning assisted convolutional neural network for heart disease prediction on the internet of medical things platform. *IEEE Access*, 8, 189503-18951.

Roth, G. A., Mensah, G. A., Johnson, C. O, Addolorato G., Ammirati E., Baddour L.M., Barengo N.C., Beaton A.Z., Benjamin E.J., Benziger C.P., Bonny A., Brauer M., Brodmann M., Cahill T.J., Carapetis J., Catapano A.L., Chugh S.S., Cooper L.T., Coresh J., Criqui M., DeCleene N., Eagle K.A., Emmons B.S., Feigin V.L., Fernández J., Fowkes G., Gakidou E., Grundy S.M., He F.J., Howard G., Hu F., Inker L., Karthikeyan G., Kassebaum N., Koroshetz W., Lavie C., Lloyd J.D., Lu H.S., Mirijello A., Temesgen A.M., Mokdad A., Moran A.E., Muntner P., Narula J., Neal B., Ntsekhe M., Moraes de Oliveira G., Otto C., Owolabi M., Pratt M., Rajagopalan S., Reitsma M., Ribeiro A.L., Rigotti N., Rodgers A., Sable C., Shakil S., Sliwa H.K., Stark B., Sundström J., Timpel P., Tleyjeh I.M., Valgimigli M., Vos T., Whelton P.K., Yacoub M., Zuhlke L., Murray C. & Fuster V.(2020). Global burden of cardiovascular diseases and risk factors, 1990–2019: update from the GBD 2019 study. *Journal of the American College of Cardiology*, 76(25),2982–3021.

Roy T. S., Roy J. K. & Mandal N. (2022). Classifier identification using deep learning and machine learning algorithms for the detection of valvular heart diseases. *Biomedical Engineering Advances*, 3, 100035.

Saikumar K., Rajesh V., Srivastava G. & Lin J. C. (2022). Heart disease detection based on internet of things data using linear quadratic discriminant analysis and a deep graph convolutional neural network. *Computational neuroscience*, 16,1-13.

Sajja T. K. & Kalluri H. K. (2020). A Deep Learning Method for Prediction of Cardiovascular Disease Using Convolutional Neural Network. *Revue d'Intelligence Artificielle*, 34(5), 601-606.

Sandhya Y. (2020). Prediction of Heart Diseases using Support Vector Machine. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 8(2), 126-135.

Sheeba A., Padmakala S., Subasini C. A. & Karuppiah S. P. (2022). MKELM: Mixed Kernel Extreme Learning Machine using BMDA optimization for web services based heart disease prediction in smart healthcare. *Computer Methods in Biomechanics and Biomedical Engineering*, 25(10), 1180-1194.

Shrivastava P. K., Sharma M., & Kumar A. (2023). HCBiLSTM: A hybrid model for predicting heart disease using CNN and BiLSTM algorithms. *Measurement: Sensors*, 25, 100657.

Simegn G. L., Gebeyehu W. B. & Degu M. Z. (2022). Computer-Aided Decision Support System for Diagnosis of Heart Diseases. *Research Reports in Clinical Cardiology*, 39-54.

Swain, D., Parmar, B., Shah, H., Gandhi A., Pradhan M.R., Kaur H. & Acharya B. (2022). Cardiovascular Disease Prediction using Various Machine Learning Algorithms. *Journal of Computer Science*, 18 (10), 993-1004.

Swathy M. & Saruladha K. (2022). A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques. *ICT Express*, 8(1), 109-116.

Tasnim F. & Habiba S. U. “A comparative study on heart disease prediction using data mining techniques and feature selection,” *International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*, 2021.

Taylor J.S. & Cristianini N. 2011. Kernel Methods for Pattern Analysis, Cambridge University Press, Kingdom of England, 47-84.

Ting D. S., Pasquale L. R., Peng, L., Campbell J.P., Lee A.Y., Raman R., Wei G.S., Schmetterer, L., Keane P.A. & Wong T.Y. (2019). Artificial intelligence and deep learning in ophthalmology. *British Journal of Ophthalmology*, 103(2), 167-175.

Uddin M. N. & Halder R. K. (2021). An ensemble method based multilayer dynamic system to predict cardiovascular disease using machine learning approach. *Informatics in Medicine Unlocked*, 24, 100584.

Umer M., Sadiq S., Karamti H., Karamti W., Majeed R. & Nappi M. (2022). IoT Based Smart Monitoring of Patients' with Acute Heart Failure. *Sensors*, 22(7), 2431.

Vaduganathan M., Mensah G.A., Turco J.V., Fuster V. & Roth, G.A. (2022). The Global Burden of Cardiovascular Diseases and Risk: A Compass for Future Health. *Journal of the American College of Cardiology*, 80(25), 2361-2371.

Wang X., Wang J., Wang W., Zhu M., Guo H., Ding J., Sun J., Zhu D., Duan Y., Chen X., Zhang P., Wu, Z. & He K. (2022). Using artificial intelligence in the development of diagnostic models of coronary artery disease with imaging markers: A scoping review. *Frontiers in Cardiovascular Medicine*, 9,1-11.

Wankhede J., Sambandam P. & Kumar, M. (2022). Effective prediction of heart disease using hybrid ensemble deep learning and tunicate swarm algorithm. *Journal of Biomolecular Structure and Dynamics*, 40(23), 13334-13345.

Waqar M., Dawood H., Dawood H., Majeed N., Banjar A. & Alharbey R. (2021). An efficient SMOTE-based deep learning model for heart attack prediction. *Scientific Programming*, 1-12.

WHO (2018). Non- Communicable Diseases Country Profile: India. https://apps.who.int/iris/bitstream/handle/10665/274512/9789241514620-eng.pdf

WHO (2021). Cardiovascular disease. https://www.who.int/health-topics/cardiovasculardiseases#tab=tab_1

WHO (2021). Cardiovascular Diseases(CVDs). World Health Organization Publishing. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

Win T. Z. & Kham N. S. (2019). Information gain measured feature selection to reduce high dimensional data, Doctoral dissertation, MERAL Portal.

Yan J., Tian J., Yang H., Han G., Liu Y., He H., Han Q. & Zhang Y. (2022). A clinical decision support system for predicting coronary artery stenosis in patients with suspected coronary heart disease. *Computers in Biology and Medicine*,151, 106300.

You Y., Wang W., Li D., Jia Y., Li D., Zeng R., & Zhang L. (2023). Multi-modal machine learning based on electrocardiogram data for prediction of patients with ischemic heart disease. *Electronics Letters*, 59(2),1-4.

Zhang A., Lipton Z. C., Li M., & Smola A. J. 2021. Dive into deep learning. arXiv preprint, NY,1222.