



A Scoping Review of Machine Learning Techniques and Their Utilisation in Predicting Heart Diseases

Maad M. Mijwil

Computer Techniques Engineering Department,
Baghdad College of Economic Sciences University,
Baghdad, Iraq.
mr.maad.alnaimiy@baghdadcollege.edu.iq

Ban Salman Shukur

Computer Science Department , Baghdad
College of Economic Sciences University,
Baghdad, Iraq.
dr_bansalman@baghdadcollege.edu.iq

Article history: Received 1 March 2022, Accepted 8 May 2022, Published in July 2022.

Doi: 10.30526/35.3.2813

Abstract

Heart diseases are diverse, common, and dangerous diseases that affect the heart's function. They appear as a result of genetic factors or unhealthy practices. Furthermore, they are the leading cause of mortalities in the world. Cardiovascular diseases seriously concern the health and activity of the heart by narrowing the arteries and reducing the amount of blood received by the heart, which leads to high blood pressure and high cholesterol. In addition, healthcare workers and physicians need intelligent technologies that help them analyze and predict based on patients' data for early detection of heart diseases to find the appropriate treatment for them because these diseases appear on the patient without pain or noticeable symptoms, which leads to severe concerns such as heart failure and stroke and kidney failure. In this regard, the authors highlight an amount of literature considered the most practical in utilizing machine learning techniques in predicting heart disease. Twenty articles were chosen out of fifty articles gathered and summarised in a table form. The main goal is to make this article a reference that can be utilized in the future to assist healthcare workers in studying these techniques with ease and saving time and effort on them. This article has concluded that machine learning techniques have a significant and influential role in analyzing disease data, predicting heart disease, and assisting decision-making. In addition, these techniques can analyze data that reaches millions of cohorts.

Keywords Artificial Intelligence, Machine Learning, Cardiology, Predication, Heart Diseases.



1. Introduction

This century is marked by the spread of harmful diseases and pandemics that claimed the lives of many people, so their number decreased, their standard of living decreased, and it has become challenging to recover from these diseases, as the patients suffer from different stages of the development of the disease in their body [1]. The most famous pandemics experienced by the inhabitants of the planet are the coronaviruses (COVID-19) [2], and the most famous diseases are heart diseases [3], which may lead to sudden death. In addition, chronic diseases are prevalent diseases that appear obviously in the elderly and lead to a change in their lifestyle and diet. The most common reason for the emergence of these diseases is smoking and alcohol [4], in addition to other causes. The most famous of these diseases are diabetes, hypertension, heart diseases, chronic obstructive pulmonary disease, arthritis (joint inflammation), cancer, asthma, and Alzheimer. Fortunately, sciences and technology have made worthwhile and significant contributions to the medical domain in modeling, predicting, and controlling diseases and pandemics that humanity is fighting against and helping clinicians and healthcare workers to make decisions [5]. Artificial intelligence is among the most famous of these sciences, which can predict and track the spread of the COVID-19 pandemic disease [6]. Artificial intelligence is the science of creating programmed machines, thinking like humans, and imitating humans [7]. In other words, these machines simulate human intelligence in every way of life [8]. It is utilized among humans to describe machines that mimic the brain's cognitive functions [9]. For instance, machines learn the logical interactions shown by humans or animals and do the same. Programs that simulate brain features such as learning, problem-solving, and a similar brain structure are programmed and designated on the machine. After learning the behavior, the machine acts similarly to solve any problem it encounters. The most helpful feature of artificial intelligence is that actions with a high probability of occurring can be prepared and applied in real life to achieve an explicit purpose. Besides, artificial intelligence has various sub-components: machine learning, artificial neural networks, computer vision, deep learning, cognitive computing, and natural language processing. Machine learning and deep learning are the most influential, widespread, and widely accustomed sub-components. The concept that computer programs automatically learn from new data without the help of humans and can adapt to it when the data is modified this technique is called "machine learning" [10]. For deep learning [11], it is machine learning by capturing and using large amounts of unstructured data such as text, images, or videos and training the model on this data.

Artificial intelligence in healthcare involves designing computer algorithms to execute assignments generally associated with human intelligence. Also, artificial intelligence is widely employed in technical and standard glossaries to encompass a range of learning, including but not limited to machine learning, feature learning, deep learning, and natural language processing. Whatever the detailed technology, the overall purpose of these techniques in medicine is to operate computer algorithms to extract relevant information from data and aid decision-making in the clinical stage. On the other hand, artificial intelligence can enhance the medical crew and patient effects through clinical decision support, patient monitoring, healthcare interventions such as patient management, aiding in diagnosis and treatment selection, risk prediction and disease stratification, reducing medical errors and increasing productivity, reducing costs, and improving community productivity as well as providing important opportunities such as the influence on health. **Figure 1** shows healthcare workers

using artificial intelligence techniques to get their work done. Personal health data is accumulated from multiple technology platforms such as web servers, electronic health records (EHRs), genetic data, personal computers, smartphones, mobile applications, wearable devices and sensors, demographic information, and healthcare. In concise, artificial intelligence can go beyond the cognitive ability of humanity to manage information effectively, as it has the role of extracting notes from x-ray images [12], laboratory results, genetic data, etc., efficiently and helping healthcare workers determine the patient's condition accurately. Machine learning techniques have developed significantly and remarkably recently. Physicians and healthcare workers have utilized them to support them in analyzing and predicting heart diseases of all types due to their excellent ability to give effects that help make decisions about the patient's condition.



Figure 1. Healthcare workers using artificial intelligence techniques [downloaded from Google].

The principal contribution of this article is to summarise the role of machine learning in predicting heart disease by focusing on the 20 most influential papers issued between 2017 and 2021. The authors try to make this an essential reference for future use. In addition, the article shows the vital role of machine learning in analyzing heart images and understanding heart diseases and how they impact human life.

This article is outlined as follows; machine learning and its techniques are detailed in Section two, followed by Section three, which shows the articles selected by the authors as they are presented in a table with their details. Finally, a conclusion is drawn up in section four.

2. Machine Learning Techniques

Artificial intelligence, machine learning, and deep learning are concepts that are often perplexed by each other. Chiefly, artificial intelligence is the ability of programs to learn and act like humans, whereas machine learning is algorithms written for the same purpose. Machine learning has noticed tremendous interest in the last few years, and this matter is still going ahead in persistent steps [13]. Nevertheless, people look at machine learning as an advanced technology that could be developed, applied, and accessed only by experts. This view is becoming less adopted, and now more categories of professionals are interested in using or adopting machine learning and other artificial intelligence methods and tools to support their investigation and work [14]. It is built on the sub-domains of mathematics such as probability, statistics, and optimization. Also, it is a science that will continue to be an ever-expanding domain. There are many motivations for this. First, separate research communities in symbolic machine learning, computational learning theory (CoLT), statistics, neural networks, and pattern recognition discovered each other and began to work together.

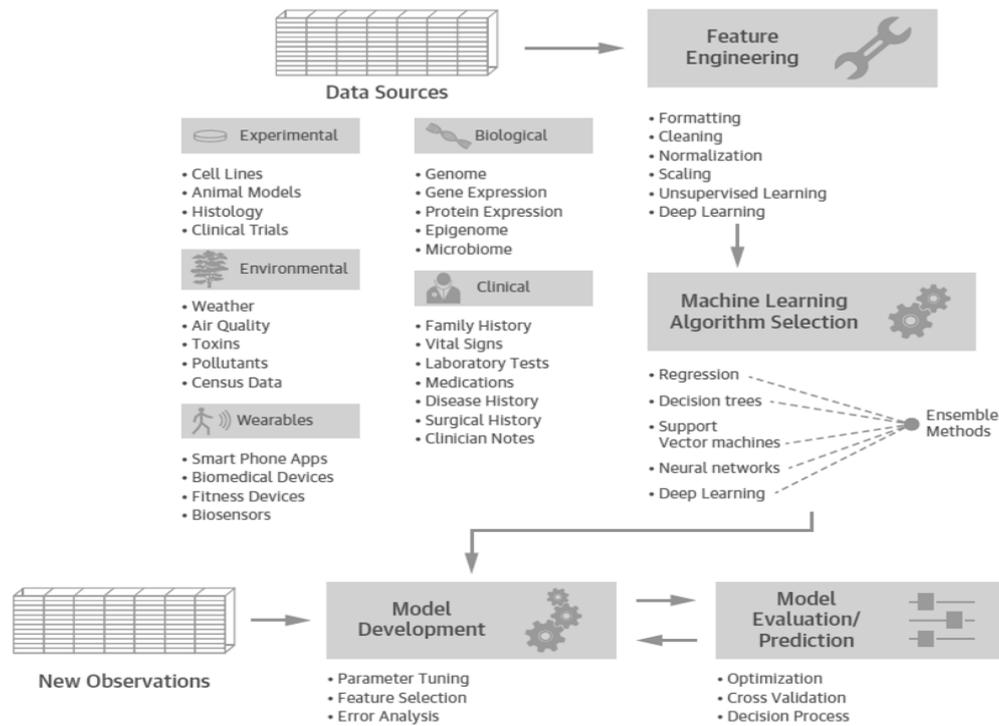


Figure 2. The Role of artificial intelligence in the healthcare domain [15].

Second, machine learning techniques can be involved in new problems such as knowledge discovery in databases, robot control, language processing, and co-optimization, as well as traditional issues such as speech recognition, handwriting recognition, face recognition, and medical problems. In healthcare, researchers are now more interested in adopting and using artificial intelligence and machine learning techniques to gain more knowledge and help (see Figure 2). For instance, machine learning techniques can be a good source of originating hidden knowledge from medical data and records to assist society in reducing the number of trials required to diagnose a person's disease accurately. A survey shows that half of the organizations classified as healthcare are using or scheduling the use of artificial intelligence in imaging. Moreover, machine learning techniques are naturally divided into supervised and unsupervised [15]. Supervised learning is based on training a data sample from the data source with the correct pre-defined classification while self-organising neural networks learn to utilise an unsupervised learning algorithm to identify hidden patterns in unlabelled input data. Unattended refers to learning and organising information without providing an error signal to evaluate a solution. The type of supervised learning is divided into classification and regression. It is possible to use different algorithms according to the dataset's content. The most influential machine learning techniques will be briefly addressed in this section.

2.1 Naïve Bayes

This classification algorithm relies on a set of influential hypotheses for the independence of covariates and makes them consistent with Bayes theory [16]. Moreover, this algorithm assumes independence between the predictor variables conditional on the response and a Gaussian distribution of numeric predictors calculated from the training dataset with mean and standard deviation. Also, this classifier is the most comfortable supervised machine learning method. It is employed by few data mining practitioners at the expense of traditional methods such as decision trees or logistic regressions [17]. In addition, the benefit of this

technique is the simplicity of programming, comfort, and speed of parameter estimation (even in massive databases). Despite its benefits, its little use in practice comes partly from the fact that there is no simple explicit model (interpretation of preconditional probability); the practical usefulness of such a technique is called into question.

2.2 Support Vector Machine

SVM algorithm is mainly made to solve classification problems but has been grown and reworked over the years. Its most famous variant is the SVM for regression, SVM for solving integral equations, SVM for estimating density support, and SVM that uses different soft-margin costs and parameters. It is considered one of the broadly used algorithms in the machine learning subject within the cardiac domain. For instance, SVM can classify data in binary problems perfectly by finding the optimal solution to distinct data points of first-class from other data acting as a second class. It is a straightforward algorithm capable of solving complex nonlinear relationships, making it very suitable to be employed in heart disease prediction systems containing patient records with binary data [18]. In [19] mentions that SVM has 90% accuracy in predicting in-stent restenosis from plasma metabolite levels. Also, this technique tries to fetch the optimal hyperplane, which maximizes the distance from the nearest training data points of any class utilized in classification problems because of having influential capabilities in generalizing new unrecognized data items, flexible non-linear decision boundaries, and their dependency on few hyper-parameters.

2.3 Logistic Regression:

This technique is one of the most expressive, versatile, and diverse techniques for analyzing clinical and pandemics. It is a statistical model to interpret relations between a set of qualitative variables and a generalized linear model that uses a logistical function as a link function. Moreover, it is utilized to predict the probability of an event occurring, such as indicating a particular disease in the human body [20].for example: A patient dies or not before discharge.

- A person stops smoking or not after treatment.
- In a retrospective study, an individual is either a case or a control.
- Whether or not an HIV-positive patient is in stage IV.

Besides, logistic regression is a worthwhile quantitative procedure for concerns where the dependent variable takes values in a finite set. It was created in the 1960s by three scholars, Confield, Gordon, and Smith [21], but its actual use began in the 1980s due to its computational facilities.

2.4 K-Nearest Neighbours

Classification of query points whose class is unknown is from the goal of the K-nearest neighbours' technique due to their respective distances to point in a learning set [22]. This technique supposes that each example in the learning set is a random vector. In short, its purpose is to classify the quantitative or qualitative dataset based on fourth votes, namely metrics, kernels, overlap metrics, and value difference metrics. On the other hand, it is an essential and straightforward classification technique.

It is explicitly used when there is little or no information about the data distribution. Also, it is a non-parametric technique [23]. This means that it does not make any presumptions about the data distribution used in the analysis. It can handle dimensionality reduction tasks in a unified

manner and is suitable for realistic environments where actual data availability doesn't follow the theoretical statistics like in normal distribution. Therefore, KNNs do not make any generalization and keep all data because it uses a quick training step.

2.5 Random Forest

It is considered one of the most widely used techniques in predicting and data analysis. We can convey that most studies employ machine learning techniques, including the random forest technique. It is a supervised machine learning technique operated on classification and regression issues. Furthermore, a random forest is an ensemble classifier [24]; that is, it consists of a large number of individual decision groups. Each Tree within its distribution provides a prediction of a specific event. The class with the most significant votes is considered the typical prediction. Each Tree in the forest collects random samples from the dataset with replacement; this process is called bagging (bootstrap aggregation) [25]. The out-of-bag score will give a complete description of the model's performance during the training phase.

2.6 Linear Regression

This technique is considered one of the most statistical techniques utilized to verify or estimate the relationship between dependent variables and a set of independent descriptive variables within a dataset [26]. In general, this technique is employed in data analysis. However, it is not influenced by the size of the database, as it has the ability in a qualitative research method to model and analyze many variables. In this technique, the dependent variable is a predictive or descriptive element and can be described as the result or response to a particular query within a large dataset. Moreover, this technique allows examining the relationship between two or more variables and identifying the most significant changes on a topic of interest, especially in medical data, because they are extensive data. Variables are divided into two styles: the dependent, which is the factor that tries to understand or predict, and the second is the independent variable, which is the factor that affects the special dependent variable.

2.7 Linear Discriminant Analysis

It is the standard feature extraction technique in pattern classification problems [27]. It is defined as a linear model for classification and dimensionality reduction. This technique is distinguished by dropping data from the dimensional feature space to the dimension space to increase the variability between the categories and reduce the variability within the categories. In short, the most crucial feature of this technique is the symmetric squared distance effects. This technique is an alternative to logistic regression when the qualitative variable has more than two levels. In addition, this technique is distinguished from logistic regression by the following:

- If the categories are sufficiently separated, then the parameters estimated in the logistic regression model are inconsistent. The LDA method does not suffer from this issue.
- If the number of observations is low and the distribution of the predictors is nearly normal in each category, the LDA will be more regular than the logistic regression.

2.8 Learning Vector Quantisation

It is a supervised machine learning method and is a type of artificial neural network inspired by biological models of neural systems [28]. Moreover, this technique can self-organize its network training and deals with the problem of multi-category classification. Also, it contains two layers of input and output. In general, LVQ is considered a model to classify learning which establishes a promising alternative for deep networks, mainly relying on Euclidean metrics to compare data vectors with prototype vectors. It is a known algorithm specialized in classifying patterns or selecting prototypes through its network that depends on nearest neighbour patterns in its classification process and is appropriate for solving non-linear separation problems and classification of large amounts of data [29]. By applying the smallest Euclidean distance, the LVQ algorithm is designed to decide which vector is winning, which will cause this vector to be chosen [30].

3. Machine Learning in Cardiology

Machine learning has the extraordinary ability to predict vital information from a large set of data [31]. Physicians can make decisions based on the effects of these techniques. For instance, the medical field is rich in information that arises from the womb of laboratory examinations and clinical and physiological observations. Physicians and healthcare workers started analyzing patient data with different data and organized algorithms that depend on constantly updated data sets to improve the ability to diagnose a disease or predict patient outcomes. In addition, machine learning cannot replace the physician, but physicians who utilize machine learning techniques will replace traditional physicians who are not keeping pace with artificial intelligence techniques. In this section, a list of contributions of machine learning techniques in diverse fields of cardiology in the last five years from 2017 to 2021 that we consider to be of paramount importance is included with a detailed description of each study in terms of the type of disease that is predicted, the number of patients and the names of the techniques employed in the prediction process with the conclusions of the study as illustrated in Table 1. To clarify, the literature is selected based on two main factors: the number of techniques applied and the accurate data about patients (Patient Cohorts).

Table 1. A bunch of literature indicates heart disease for patient cohorts utilising machine learning techniques between 2017-2021.

References	Disease type	Utilisation	Patient Cohorts	Techniques	Outcomes
Weng <i>et al.</i> [32]	Cardiovascular Risk	Prediction heart disease over ten years, where data collected are free of cardiovascular disease and likened with American College of Cardiology guidelines.	378,256 patients from UK family practices	RF, LR, GB, and NN	The highest executed model is NN with an AUC of 0.764.

Budzianowski et al. [33]	Arrhythmia	Prediction of early recurrence of atrial fibrillation following cryoballoon ablation	118 patients with 56 clinical signs	Over-Sampling, SVM and GBT	Techniques have achieved in predicting a range of early matters of the disease
Nanayakkara et al. [34]	Heart failure	Predict the probability of dying from cardiac arrest in hospital	39,566 patients from Australian and New Zealand	LR, GBM, SVC, RF, ANN, and Ensemble	The most useful technique in predicted mortality is ANN with an accuracy of 46.7% and AUC 0.85.
Yıldırım et al. [35]	Arrhythmia	Prediction cardiac arrhythmia based on 17 classes of MIT - BIH Arrhythmia database and employing the effects of long-duration electrocardiography (ECG) signal analysis.	based on 1000 ECG signal from 45 patients	1D-CNN	The accuracy is 91.33% with time per single sample of 0.015 s
Ul Haq et al. [36]	Heart failure	Prediction heart failure based on patient diagnostic data through a hybrid intelligent system framework based on seven machine learning classifiers and three feature selection algorithms.	303 patients	LR, KNN, ANN, SVM, NB, and DT	The most acceptable classifier of prediction is SVM. [Accuracy=86%, Specificity=88%, Sensitivity=78%, and AUC=86%]
Tesche et al. [37]	Coronary CT Angiography	Predicting heart disease by studying coronary CT angiography-derived fractional flow reserve	85 patients	MLA	Sensitivity of 79% & Specificity of 93%
Attia et al. [38]	Heart failure	Prediction atrial fibrillation during sinus rhythm for patients aged 18 and over, based on	180,922 patients	CNN	AUC of 0.87, sensitivity of 79%, specificity of 79.5%, F1 score of 39.2%,

		laboratory data between 1993 and 2017.			and accuracy of 79.4%.
Dinh et al. [39]	Cardiovascular Risk	Predicting diabetes and cardiovascular disease	131 variables	LR, SVM, RF, and GB	The most beneficial effect is GB earned an AUC of 86.2% (without laboratory data) and 95.7% (with laboratory data).
Oikonomou et al. [40]	Coronary CT Angiography	Prediction medical risks employing coronary CT angiography to detect perivascular fat by relying on three types of patients	312 patients in three cases	RF	The result of RF is C-statistic 0.77 [95% CI: 0.62–0.93]
Hill et al. [41]	Arrhythmia	Predicting a severe heart disease, atrial fibrillation, which is one of the most common heart diseases for cohort adults whose age ≥ 30 years without a history of atrial fibrillation	2,994,837 patients	NN, LR, RF, SVM, and CR	The best result is LR: 75% sensitivity and AUC= 0.695
Tiwari et al. [42]	Arrhythmia	Prediction of incident atrial fibrillation for patients with arrhythmia depending on the electronic health history data from Jan 1, 2011, to Oct 1, 2018	2,252,219 patients from three large hospitals in Colorado, USA.	NB, LR, RF, NN, and GBM	The most acceptable prediction effect is for NN: F1-score is 0.110 and AUC is 0.800
Alhusseini et al. [43]	Atrial Fibrillation	To Classify Intracardiac Electrical Patterns During Atrial Fibrillation based on Hilbert-transform to produce 175 000	35 patients	CNN	The accuracy is 95%

image grids in patients					
Loring <i>et al.</i> [44]	Atrial Fibrillation	Prediction of the effect of atrial fibrillation by analysing the ORBIT-AF and GARFIELD-AF registries to determine the death and stroke	74,792 patients	LR, RF, GB, and NNs	The most helpful performance is provided by the LR model with results of predications as follow: for death [AUC = 0.80 in ORBIT-AF, 0.75 in GARFIELD-AF] and stroke [AUC = 0.67 in ORBIT-AF, 0.66 in GARFIELD-AF]
Ward <i>et al.</i> [45]	Atherosclerotic cardiovascular disease	Prediction atherosclerotic risk of cardiovascular disease in multi-ethnic folks	797,505 patients	LR, RF, GBM, and EGB	The most suitable performing model in the test is GBM (AUC 0.835, 95% CI: 0.825–0.846)
Krittanawong <i>et al.</i> [46]	Coronary artery disease, Heart failure, and Arrhythmia	Prediction and investigation of heart diseases and stroke	3,377,318 patients	BM, SVM, and CNN	AUC of BM is 0.88, AUC of SVM is 0.92, and AUC of CNN is 0.90
Vinter <i>et al.</i> [47]	Atrial Fibrillation	Predicting heart disease and distinguishing between disease in men and women based on the heart's electrical system within three months	1122 patients [female =332 & male =790]	LR and RF	Discrimination is fair for both RF (0.59 for women and 0.56 for men) and LR (0.60 for women and 0.59 for men).
Wang <i>et al.</i> [48]	Arrhythmia	Expect the happening of arrhythmia after acute myocardial infarction	2084 patients with acute myocardial infarction	DT, RF, and ANN	The optimal execution is ANN with an accuracy of 0.668 and AUC of 0.654

Lip <i>et al.</i> [49]	Atrial Fibrillation	Prediction of atrial fibrillation in older adults with COVID-19 and non-infected.	280,592 patients from USA	LR	The effect of LR is 0.729 (95%CI 0.718-0.740)
Wang <i>et al.</i> [50]	Heart failure	Prediction of malignant arrhythmia in hospitalized patients with heart failure	2794 hospitalized patients	LLR-1, LLR-2, MARS, CART, RF, and XGBoost	The best effect gained by this investigation is the XGBoost technique with 0.998 AUC
Khurshid <i>et al.</i> [51]	Atrial Fibrillation	The prediction is based on ECG data from three hospitals to infer 5-year happening AF risk employing ECGs in patients obtaining longitudinal primary care.	45,770 patients (53% women and 47% men) from Brigham and Women's Hospital and UK Biobank	CNN	The findings indicate that CNN utilising ECG to estimate AF risk is robust and valid [AUC 0.823 (0.790-0.856) for 5-year] for patients at Massachusetts General Hospital

5. Conclusion

More often than not, machine learning for some physicians is complicated mathematics, meaning that their responses are not optimistic when they hear the pronunciation of the word machine learning. There are also appropriate responses and a great desire to utilize machine learning techniques. These technologies have a fantastic future in making predictions in determining heart diseases and require cardiologists to interact and cooperate fully with the practices of these technologies. For this collaboration to grow, cardiologists need a high-level knowledge of the role presented by these technologies, awareness of machine learning efforts in detecting and analysing heart disease, whether from patient data or through images, and knowledge of potential shortcomings in machine learning techniques. Most studies that have been issued use machine learning techniques in analyzing heart images and other assignments.

Nevertheless, these techniques have a significant role in extracting patient data and predicting disease regardless of the large number of patient data or the increase in the size and variety of the training data set. Despite the significant effect of these techniques, no published clinical trials have compared these techniques to the human evaluation of the data set. Therefore, it is required to achieve future controlled clinical trials to demonstrate the proficiency and efficacy of these techniques in clinical practice. In addition, confirmation should be carried out not using data from the same group employed in training but also from other groups so that these techniques can be analyzed more efficiently. In the future, a group of studies will be executed on the use of machine learning techniques in analysing heart data and predicting potential human diseases.

Nomenclatures

Acronym	Description
ANN	Artificial Neural Network
AUC	Area Under the Curve
CART	Classification and Regression Trees
CNN	Convolutional Neural Network
COVID-19	Coronavirus 2019
CR	Cox Regression
DT	Decision Trees
EGB	Extreme Gradient Boosting
GB or GBT	Gradient Boosting
GBM	Gradient Boosting Machine
KNN	K-Nearest Neighbours
LR	Logistic Regression
MARS	Multivariate Adaptive Regression Splines
MLA	Machine Learning Algorithm
NB	Naïve Bayes
NN	Neural Network
RF	Random Forest
SVM	Support Vector Machine

References

- Hoffman, D. J.; Reynolds, R. M.; Hardy, D. B. Developmental origins of health and disease: current knowledge and potential mechanisms. *Nutr Rev.* **2017**, *75* (12), 951-970.
- Mijwil, M. M.; Al-Mistarehi, AH; Muttur, D. S. The Practices of Artificial Intelligence Techniques and Their Worth in the Confrontation of COVID-19 Pandemic: A Literature Review. *Mob Forensics J.* **2022**, *4*(1), 11-30.
- Núñez-Gil, I. J.; Fernández-Ortiz, A.; Eid, C. M.; Huang, J.; Romero, R.; Becerra-Muñoz, V. M. Underlying heart diseases and acute COVID-19 outcomes. *Cardiol J.* **2020**, *28* (2), 202-214.
- Taylor, K.; Elhakeem, A.; Nader, J. L. T.; Yang, T. C.; Isaevska, E.; Richiardi, L.; Vrijkotte, T.; Moira, A. P.; Murray, D. M.; Finn, D.; Mason, D.; Wright, J.; Oddie, S.; Roeleveld, N.; Harris, J. R.; Andersen, A. N.; Caputo, M.; Lawlor, D. A. Effect of Maternal Prepregnancy/Early-Pregnancy Body Mass Index and Pregnancy Smoking and Alcohol on Congenital Heart Diseases: A Parental Negative Control Study. *J Am Heart Assoc.* **2021**, *10*(11), 1-78.
- Mijwil, M. M.; Al-Mistarehi, AH.; Aggarwal, K. The Effectiveness of Utilising Modern Artificial Intelligence Techniques and Initiatives to Combat COVID-19 in South Korea: A Narrative Review. *Asian J Appl Sci.* **2021**, *9* (5), 343-352.
- Aggarwal, K.; Mijwil, M. M.; Al-Mistarehi, AH.; Alomari, S.; Gök, M.; Alaabdin, A. M.; Abdulrhman, S. H. Has the Future Started? The Current Growth of Artificial Intelligence, Machine Learning, and Deep Learning. *Iraqi J Comput Sci Math.* **2022**, *3*(1), 115-123.

7. Ostheimer, J.; Chowdhury, S.; Sarfraz, I. An alliance of humans and machines for machine learning: Hybrid intelligent systems and their design principles. *Technol Soc.* **2021**, *66*, 101647.
8. Guzman, A. L.; Lewis, S. C. Artificial intelligence and communication: A Human–Machine Communication research agenda. *New Media Soc.* **2019**, *22(1)*, 70-86.
9. Mijwil, M. M.; Salem, I. E.; Abttan, R. A. Utilisation of Machine Learning Techniques in Testing and Training of Different Medical Datasets. *Asian J Comput Inf Sys.* **2021**, *9(5)*, 29-34.
10. Matsuzaka, Y.; Uesawa, Y. A Molecular Image-Based Novel Quantitative Structure-Activity Relationship Approach, DeepSnap-Deep Learning and Machine Learning. *Curr Issues Mol Biol.* **2020**, *42(1)*, 455-472.
11. Rasheed, J.; Hameed, A. A.; Djeddi, C.; Jamil, A.; Al-Turjman, F. A machine learning-based framework for diagnosis of COVID-19 from chest X-ray images. *Interdiscip Sci.* **2021**, *13*, 103-117.
12. Mijwil, M. M. Implementation of Machine Learning Techniques for the Classification of Lung X-Ray Images Used to Detect COVID-19 in Humans. *Iraqi J Sci.* **2021**, *62 (6)*, 2099-2109.
13. Farhan, B. I.; Jasim, A. D. A Survey of Intrusion Detection Using Deep Learning in Internet of Things. *Iraqi J Comput Sci Math*, **2022**, *3 (1)*, 83-93.
14. Rudin, C.; Ustun, B. Optimized Scoring Systems: Toward Trust in Machine Learning for Healthcare and Criminal Justice. *INFORMS J Appl Anal.* **2018**, *48 (5)*, 449-466.
15. Johnson, K. W.; Soto, J. T.; Glicksberg, B. S.; Shameer, K.; Miotto, R.; MPhil, M. A.; Ashley, E.; Dudley J. T. Artificial Intelligence in Cardiology. *J Am Coll Cardiol.* **2018**, *71(23)*, 2668-2679.
16. Rimal, Y. Naïve Bayes Machine Learning Classification with R Programming: a Case Study of Binary Data Sets. *Int J Orange technol.* **2019**, *1 (2)*, 27-34.
17. Albashrawi, M. Detecting Financial Fraud Using Data Mining Techniques: A Decade Review from 2004 to 2015. *J Data Sci.* **2016**, *14(3)*, 553-569.
18. Djerioui, M.; Brik, Y.; Ladjal, M.; Attallah, B. Neighborhood Component Analysis and Support Vector Machines for Heart Disease Prediction. *Ing des Syst d'Information.* **2019**, *24 (6)*, 591-595.
19. Cui, S.; Li, K.; Ang, L.; Liu, J.; Cui, L.; Song, X.; Lv, S.; Mahmud, E. Plasma Phospholipids and Sphingolipids Identify Stent Restenosis After Percutaneous Coronary Intervention. *JACC Cardiovasc Interv.* **2017**, *10 (13)*, 1307-1316.
20. Laitinen, E. K.; Laitinen, T. Bankruptcy prediction: Application of the Taylor's expansion in logistic regression. *Int Rev Financ Anal.* **2020**, *9 (4)*, 327-349.
21. Pérez, M. D. F.; Pérez, J. C. R. La regresión logística: una herramienta versátil. *Nephrology.* **2000**, *20, (6)*, 477-565.
22. Rodrigues, É. O. Combining Minkowski and Chebyshev: New distance proposal and survey of distance metrics using k-nearest neighbours classifier. *Pattern Recognit Lett.* **2018**, *110*, 66-71.
23. Abedi, R.; Bonyad, A. E.; Moridani, A. Y.; Shahbahrami, A. Evaluation of IRS and Landsat 8 OLI imagery data for estimation forest attributes using k nearest neighbour non-parametric method. *Int J Image Data Fusion.* **2018**, *9(4)*, 287-301.

24. Belgiu, M; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J Photogramm Remote Sens.* **2016**, *114*, 24-31,.
25. Wang, Y.; Xiu, C.; Zhang, X.; Yang, D. WiFi Indoor Localization with CSI Fingerprinting-Based Random Forest. *Sensors*, **2018**, *18* (9), 1-23.
26. Ghosal, S.; Sengupta, S.; Majumder, M.; Sinha B. Linear Regression Analysis to predict the number of deaths in India due to SARS-CoV-2 at 6 weeks from day 0 (100 cases - March 14th 2020). *Diabetes Metab Syndr.* **2020**, *14* (4), 311-315.
27. Zhang, D.; Zhao, Y.; Du, M.; A Novel Supervised Feature Extraction Algorithm: Enhanced Within-class Linear Discriminant Analysis. *Int J Comput Sci Eng.* **2016**, *13*(1),13-23.
28. Mouloudi, S.; Rahmanpanah, H.; Gohari, S.; Burvill, C.; Tse, K. M.; Davies H. M. S. What can artificial intelligence and machine learning tell us? A review of applications to equine biomechanical research. *J Mech Behav Biomed Mater.* **2021**, *123*, 104728.
29. Badbe, V. C.; Londhe, V.; Shirole, G. Analysis of Heart Disease By LVQ in Neural Network. *Int j recent innov trends comput commun.* **2016**, *4*(4), 603-604.
30. Putri, N. K.; Rustam, Z.; Sarwinda, D. Learning Vector Quantization for Diabetes Data Classification with Chi-Square Feature Selection. *IOP Conf Ser Mater Sci Eng.* **2019**, *546*, 1-8.
31. Bharti, R.; Khamparia, A.; Shabaz, M.; Dhiman, G.; Pande, S.; Singh, P. Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning. *Comput Intell Neurosci.* **2021**, *8387680*, 1-11.
32. Weng, S. F.; Reys, J.; Kai, J.; Garibaldi, J. M.; Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data?. *PLoS ONE*, **2017**, *12*(4),1-14.
33. Budzianowski, J.; Hiczkiewicz, J.; Burchardt, P.; Pieszko, K.; Rzeźniczak, J.; Budzianowski, P.; Korybalska, K. Predictors of atrial fibrillation early recurrence following cryoballoon ablation of pulmonary veins using statistical assessment and machine learning algorithms. *Heart Vessels.* **2018**, *34*, 352-359.
34. Nanayakkara, S.; Fogarty, S.; Tremeer, M.; Ross, K.; Richards, B.; Bergmeir, C.; Xu, S.; Stub, D.; Smith, K.; Tacey, M.; Liew, D.; Pilcher, D.; Kaye, D. M. Characterising risk of in-hospital mortality following cardiac arrest using machine learning: A retrospective international registry study. *PLoS Medicine.* **2018**, *15* (11), 1-16.
35. Yıldırım, Ö.; Pławiak, P.; Tan, R.; Acharya, U. R. Arrhythmia detection using deep convolutional neural network with long duration ECG signals. *Comput Biol Med.* **2018**, *102*, 411-420.
36. Ul Haq, A.; Li, J. P.; Memon, M. H.; Nazir, S.; Sun R. A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms. *Mob Inf Syst.* **2018**, *2018*, 1-22.
37. Tesche, C.; De Cecco, C. N.; Baumann, S.; Renker, M.; McLaurin, T. W.; Duguay, T. M.; Bayer, R. R.; Steinberg, D. H.; Grant, K. L.; Canstein, C.; Schwemmer, C.; Schoebinger, M.; Itu, L. M.; Rapaka, S.; Sharma, P.; Schoepf, U. J. Coronary CT Angiography-derived Fractional Flow Reserve: Machine Learning Algorithm versus Computational Fluid Dynamics Modeling. *Radiology*, **2018**, *288*(1),64-72.
38. Attia, Z. I.; Noseworthy, P. A.; Lopez-Jimenez, F.; Asirvatham, S. J.; Deshmukh, A. J.; Gersh B. J. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. *Lancet.* **2019**, *394*(10201), 861-867.

39. Dinh, A.; Miertschin, S.; Young, A.; Mohanty, S. D. A data-driven approach to predicting diabetes and cardiovascular disease with machine learning. *BMC Medical Inform Decis Mak.* **2019**, *19*, 1-15.
40. Oikonomou, E. K.; Williams, M. C.; Kotanidis, C. P.; Desai, M. Y.; Marwan, M.; Antonopoulos, A. S. A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CT angiography. *Eur Heart J.* **2019**, *40* (43), 3529–3543.
41. Hill, N. R.; Ayoubkhani, D.; McEwan, P.; Sugrue, D. M.; Farooqui, U.; Lister, S.; Lumley, M.; Bakhai, A.; Cohen, A. T.; O'Neill, M.; Clifton, D.; Gordon, J. Predicting atrial fibrillation in primary care using machine learning. *PLoS ONE*, **2019**, *14*(11), 1-13.
42. Tiwari, P; Colborn, K. L.; Smith, D. E.; Xing, F.; Ghosh, D.; Rosenberg, M. A. Assessment of a Machine Learning Model Applied to Harmonized Electronic Health Record Data for the Prediction of Incident Atrial Fibrillation. *JAMA Netw Open.* **2020**, *3*(1),1-12.
43. Alhousseini, M. I.; Abuzaid, F.; Rogers, A. J.; Zaman, J. A. B.; Baykaner, T.; Clopton, P.; et al. Machine Learning to Classify Intracardiac Electrical Patterns During Atrial Fibrillation. *Circ Arrhythm Electrophysiol.* **2020**, *13*(8), e008160.
44. Loring, Z.; Mehrotra, S.; Piccini, J. P.; Camm, J.; Carlson, D.; Fonarow, G. C.; Fox, K. A. A.; Peterson, E. D.; Pieper, K.; Kakkar, A. K. Machine learning does not improve upon traditional regression in predicting outcomes in atrial fibrillation: an analysis of the ORBIT-AF and GARFIELD-AF registries. *EP Europace.* **2020**, *22*(11), 1635–1644.
45. Ward, A.; Sarraju, A.; Chung, S.; Li, J.; Harrington, R.; Heidenreich, P.; Palaniappan, L.; Scheinker, D.; Rodriguez, F. Machine learning and atherosclerotic cardiovascular disease risk prediction in a multi-ethnic population. *NPJ Digit Med.* **2020**, *3*(125), 1-7.
46. Krittanawong, C.; Virk, H. U.; Bangalore, S.; Wang, Z.; Johnson, K. W.; Pinotti R.; Zhang, H.; Kaplin, S.; Narasimhan, B.; Kitai, T.; Baber, U.; Halperin, J. L.; Tang, W. H. W. Machine learning prediction in cardiovascular diseases: a meta-analysis. *Sci Rep.* **2020**, *10*, 16057, 1-11.
47. Vinter, N.; Frederiksen, A. S.; Albertsen, A. E.; Lip, G. Y. H.; Fenger-Grøn, M.; Trinquart, L.; Frost, L.; Møller, D. S. Role for machine learning in sex-specific prediction of successful electrical cardioversion in atrial fibrillation?. *OpenHeart.* **2020**, *7*, 1-7.
48. Wang, S.; Li, J.; Sun, L.; Cai, J.; Wang, S.; Zeng, L.; Sun, S. Application of machine learning to predict the occurrence of arrhythmia after acute myocardial infarction. *BMC Medical Inform Decis Mak.* **2021**, *21*(301), 1-14.
49. Lip, G. Y. H.; Genaidy, A.; Tran, G.; Marroquin, P. Estes C. Incident atrial fibrillation and its risk prediction in patients developing COVID-19: A machine learning based algorithm approach. *Eur J Intern Med.* **2021**, *91*, 53-58.
50. Wang, Q.; Li, B.; Chen, K.; Yu, F.; Su, H.; Hu, K.; et al. Machine learning-based risk prediction of malignant arrhythmia in hospitalized patients with heart failure. *ESC Heart Fail.* **2021**, *8*(6), 5363–5371.
51. Khurshid, S.; Friedman, S.; Reeder, C.; Achille, P. D.; Diamant, N.; Singh, P.; Harrington, L. X.; Wang, X.; Al-Alusi, M. A.; Sarma, G.; Foulkes, A. S.; Ellinor, P. T.; Anderson, C. D.; Ho, J. E.; Philippakis, A. A.; Batra, P.; Lubitz, S. A. ECG-Based Deep Learning and Clinical Risk Factors to Predict Atrial Fibrillation. *Circulation.* **2021**, *145*(2), 122-133.