Involving machine learning techniques in heart disease diagnosis: a performance analysis

Ban Salman Shukur[1], Maad M. Mijwil[2]
[1]Computer Science Department, Baghdad College of Economic Sciences University, Baghdad, Iraq
[2]Computer Techniques Engineering Department, Baghdad College of Economic Sciences University, Baghdad, Iraq

Article Info

Article history:

Received Apr 30, 2022
Revised Sep 19, 2022
Accepted Oct 14, 2022

Keywords:

Artificial intelligence
Cleveland clinic
COVID-19
Deep learning
Heart diseases
Machine learning

ABSTRACT

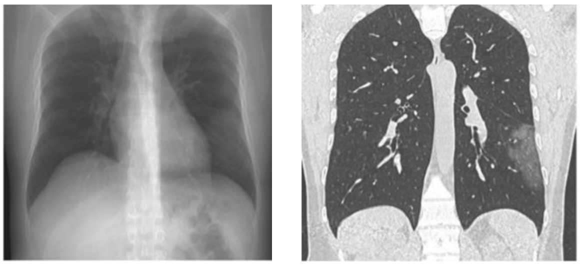
Artificial intelligence is a science that is growing at a tremendous speed every day and has become an essential part of many domains, including the medical domain. Therefore, countless artificial intelligence applications can be seen in the medical domain at various levels, which are employed to enhance early diagnosis and prediction and reduce the risks associated with many diseases, including heart diseases. In this article, machine learning techniques (logistic regression, random forest, artificial neural network, support vector machines, and k-nearest neighbors) are utilized to diagnose heart disease from the Cleveland Clinic dataset got from the University of California Irvine machine learning (UCL) repository and Kaggle platform then create a comparison between the performance of these techniques. In addition, some literature related to machine learning and deep learning techniques that aim to provide reasonable solutions in monitoring, detecting, diagnosing, and predicting heart disease and how these technologies assist in making health decisions are reviewed. Ten studies are selected and summarized by the authors published between 2017 and 2022 are illustrated. After executing a series of tests, it is seen that the most profitable performance in diagnosing heart disease is the support vector machines, with a diagnostic accuracy of 96%. This article has concluded that these techniques play a significant and influential role in assisting physicians and health care workers in analyzing heart patients' data, making health decisions, and saving patients' lives.

Corresponding Author:

Maad M. Mijwil
Computer Techniques Engineering Department, Baghdad College of Economic Sciences University
Baghdad Province, Yarmouk, Nafaq Al-Shurta, Iraq
Email: mr.maad.alnaimiy@baghdadcollege.edu.iq

1. INTRODUCTION

Today, computers and computer systems have become a vital part of our lives [1]. Computers are employed in almost every domain [2]. In the past years, computers have been benefited from summarizing large amounts of data over time and making comments about events utilizing these data while performing calculations or transmitting data [3], [4]. Nowadays, computers can make decisions about events and know the relationship between events [5]. Also, computers can also solve issues that cannot be mathematically formulated. The most noteworthy part of computer science is artificial intelligence [6], [7]. This science is distinguished by its ability to suggest computational models of learning based on human biological neural networks [8]. In other words, several models of artificial intelligence have been suggested, which, thanks to advancements in computational technology, have allowed the growth of "intelligent" techniques that facilitate processing more data in less time, speeding up the decision-making process [9], [10]. In addition, this science is involved in many domains such as advanced robotics, computer vision, natural language processing, virus tracking, prediction, diagnosis, planning and optimization. The most noteworthy part of artificial intelligence is machine learning [11]–[14], which includes different types of techniques such as random forest, naïve bayes, and k-nearest neighbor. Besides, these techniques are separated into supervised, unsupervised, and semi-supervised techniques. Moreover, machine learning is characterized by analyzing and processing big data [15], [16]. Another part of artificial intelligence is deep learning [17], [18], which is widely utilized to evaluate image, video and audio data using convolutional neural networks with their various parameters. After the spread of the coronavirus disease (COVID-19) pandemic and the development of its genes, rapid solutions are needed to monitor, detect, and diagnose diseases resulting from its causes [19]; artificial intelligence proposes many options to monitor the spread of the disease, tracking the infected and studying the behavior of the virus. Artificial intelligence techniques seek to detect infection with COVID-19 through chest X-ray or computerized tomography, see Figures 1(a) and 1(b) and the ability of these techniques to diagnose and predict heart diseases by relying on a patient dataset that is collected and analyzed to make the right decision regarding the patient’s condition [20]–[23].



(a) (b)

Figure 1. High-definition images of a 41-year-old male patient with unilateral COVID-19 pneumonia (a) X-ray and (b) computed tomography [21]

The principal contribution of this article is to execute five machine learning techniques and study their behavior in diagnosing heart disease datasets gathered from the Kaggle platform. After that, the obtained results are compared, and the most useful and worst performances among the applied techniques are known. In addition, a summary of the ten most crucial pieces of literature published between 2017 and 2022 on the use of machine learning and deep learning techniques in the diagnosis and analysis of heart disease data and what are the results of these literature so that this article can be a reference for future use.

The rest of this work is organized: in section 2, a summary of heart diseases and pieces of literature in which artificial intelligence techniques have been applied to analyze these diseases are introduced briefly. Analysis of the performance of the involved techniques with details of the results obtained in section 3. Finally, section 4 is the conclusion of this work.

2. LITERATURE REVIEW

Heart disease is a common and harmful disease that affects humans, and it is a disease caused by a change in the blood vessels of the heart or coronary arteries [24]–[27]. These diseases may occur due to other conditions, such as changes in the lung that increase pressure on the right hemisphere and lead to high blood pressure that can strain the left hemisphere [28], [29]. Changes in the balance of hormones can also cause these diseases and inflammatory rheumatic diseases [30], [31]. There are many reasons for the occurrence of heart disease that may lead to death, and therefore it may require resorting to taking appropriate treatment, according to the patient's condition. In general, heart disease occurs significantly with age, but it is very rare in children, young, and adults [32], [33]. Especially after the age of forty, in these age groups, heart diseases are prevalent due to atherosclerosis, i.e., ischemic heart disease [34]. In the recent period, the number of cases of heart disease has grown [35], especially ischemic heart disease, which causes death and is a widespread disease of death in all industrialized nations. This may be due to the high life expectancy of the population, and the incidence of ischemic heart disease increased for people aged 40 to 50 years, as illustrated in Figure 2 [36]. However, there are other dangerous factors, such as infection with the COVID-19 virus, which leads to an increase in the effort of the heart due to a lack of oxygen in the lungs [37], [38]. Today, artificial intelligence techniques are constantly striving to save patients' lives and help healthcare workers make the right decision about a patient's condition [39], [40]. Artificial intelligence has entered many domains, including healthcare, where its techniques are developing with the development of diseases, especially with heart diseases [41], [42]. In this section, a set of studies are reviewed in which these techniques are utilized to analyze heart diseases. Image analysis is an area in which machine learning progresses is faster and sufficient, therefore it is of great significance in cardiology, see Figure 3 [43]. In a study executed by Damen *et al.* [44], they suggest using machine learning techniques to analyze 4D ultrasound (4DUS) cardiac data (160,000 pixels per image with about five thousand total images) to predict left ventricular wall motility. In this work, three main features are used to build the proposed models: the projection of raw US data, smoothing spline basis across time, and parameterization of the left ventricular boundaries. They use the Monte Carlo cross-validation method to evaluate the execution of the three models. This study finds that the Monte Carlo simulation of the endocardial wall makes much closer predictions when employing Model 2 vs. Model 1 with a rate of 48.67% and employing Model 3 vs Model 2 with a rate of 83.50%. In a study conducted by Arabasadi *et al.* [45], they proposed a hybrid method for coronary artery disease diagnosis by improving the execution of the artificial neural network technique employing a genetic algorithm based on the Z-Alizadeh Sani dataset. This study found a diagnostic accuracy of more than 93%.

Figure 2. Old people suffering from heart disease in hospital [46]

A study instructed by Ankenbrand *et al.* [47] they advise using the open-source Python library (*misas*) for sensitivity analysis with arbitrary models and data. This application is applied in two cases in a cardiac magnetic resonance imaging study. This paper concludes that sensitivity analysis is a very sound tool for clinicians as it helps them interpret segmentation models. Also, Neural networks are used in decision-making, and this paper has got excellent results in analyzing heart images. Diagnosing heart diseases is one of the most essential tasks that the doctor performs to know the patient's condition, in a study conducted by Pan *et al.* [48] on the uses of machine learning techniques in diagnosing heart disease with the help of the internet of medical things. In this work, deep neural network (DNN), enhanced deep learning assisted convolutional neural network (EDCNN) techniques, neural network ensemble method (NNE), recurrent neural network (RNN), artificial neural network (ANN), and ensemble deep learning-based smart healthcare system (EDL-SHS) and are applied. The authors are able to determine the level of risk of heart disease with high adequateness, as the precision of the effects gained more than 99%. Ricciardi *et al.* [49], they apply machine learning techniques (random forests, multivariate logistic regression, ADA-Boost, and gradient boosting) in computed tomography (CT) analysis of heart disease in older adults with cardiovascular disease (CVD), coronary heart disease (CHD), and chronic heart failure (CHF).

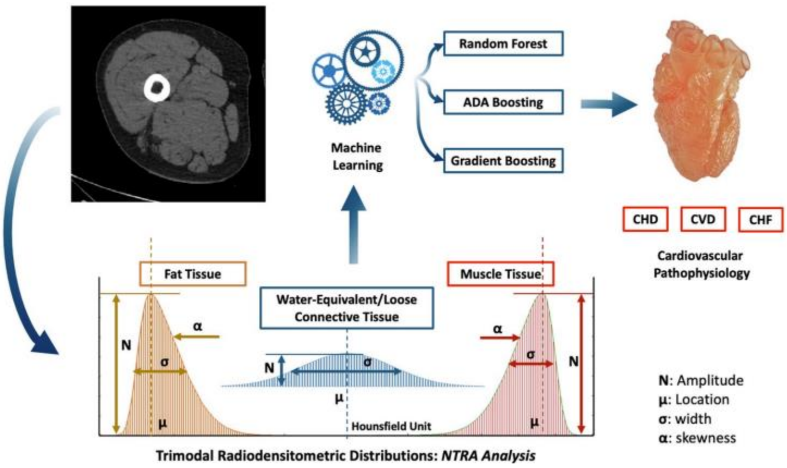


Figure 3. Heart disease images analysis mechanism utilizing machine learning techniques [49]

The effects of this study depend on four classification metrics: classification by tissue type, age, feature importance, and total classification score. The effects result in the superiority of the random forest technique as it achieves the highest classification performance for all analyses, and the overall classification scores for all three conditions as they are perfect: The AUCs for CHD=0.936, for CVD=0.914, and for CHF=0.994. In a study conducted by Helwan and Ozsahin [50], artificial neural networks are employed to recognize the left ventricle and its detection in magnetic resonance imaging of the heart. Through this analysis, a backpropagation neural network is trained on instances of the left ventricle as well as samples of the non-left ventricle. A neural network trained in backpropagation has demonstrated the power of validating simulated samples with a test set. The training and test groups gain a high accuracy rate of 100% and 88%, respectively. Many platforms support the execution of image analysis of heart disease, such as TensorFlow, Pytorch, Keras, and Caffe, that can be trained and used due to their high efficiency in the implementation. Moreover, there must be a set of images reported manually and need a great deal of experience in making the appropriate adjustments in the analysis of these images.

Alarsan and Younes [51] applied machine learning techniques (decision tree, random forests, gradient-boosted trees) and big data tool (Spark–Scala) to classify Electrocardiogram (ECG) signals based on ECG features to reveal a patient's condition. The databases of this analysis are more than 205,000 records of 51 patients. Therefore, to evaluate the execution of techniques by understanding arrhythmia. This analysis concluded that the best performance in binary classification is the gradient-boosted trees technique, which achieves an accuracy of more than 97%. In the multi-class classification, the most suitable performance is for the random forest technique, which has earned more than 98% accuracy. Ali *et al.* [52] have involved machine learning techniques (decision tree, nearest neighbor, and random forests) to diagnose early-stage heart disease by developing accurate and effective predictions based on digital patient records collected from the Kaggle platform. The performance of the techniques is compared to the accuracy of disease prediction. This analysis has found that the random forests technique has the most suitable performance, with an accuracy of 100% and a sensitivity of 100%. The authors confirm that these techniques can be used in making clinical decisions. A study by Joloudari *et al.* [53] developed a neural network and deep neural network techniques for diagnosing cardiac magnetic resonance imaging (CMRI) dataset. In this study, several operations are performed to improve the performance of both techniques in the diagnostic process. Through experiments, it is found that the best performance is the deep neural network technique, which achieved an accuracy of more than 99%, which is a very impressive effect. In comparison, the standard neural network technique achieved more than 92% accuracy.

3. PERFORMANCE ANALYSIS

In this section, five techniques (logistic regression, random forest, artificial neural network, support vector machines, and k-nearest neighbors) for diagnosing heart diseases and comparing performance are utilized to identify the most acceptable technique in terms of performance. This study relies on cardiology data from the UCL repository, including the Cleveland Clinic dataset. This dataset includes 76 parameters with a numerical value with 14 input attributes. Attribute information are sex, age, resting blood pressure, chest pain type, serum cholesterol, fasting blood sugar $> 120\, mg/dl$, resting electrocardiographic results, maximum heart rate achieved, exercise induced angina, the slope of the peak exercise ST-segment, ST depression induced by exercise relative to rest, number of major vessels (0-3) colored by fluoroscopy, num is diagnosis of heart disease (angiographic disease status): value 0:$<$ 50% diameter narrowing and value 1:$>$ 50% diameter narrowing, and thal: 3=normal; 6=fixed defect; 7=reversable defect. Figure 4 illustrates the distribution of some variables. The total data of the Cleveland Clinic is 303, which is diverged into 165 data for patients and 138 data for healthy people obtained from the Kaggle platform [54]. This work is performed utilizing Jupiter Notebook, a web-based platform for open-source application development. Moreover, it is preferable to use the Python programming language in applying machine learning techniques. The libraries Numpy, Scipy, Matplotlib, Sklearn, and Pandas are utilized during the working stages.

The logistic regression technique resulted in 23 false diagnoses and 280 proper diagnoses. The random forest technique resulted in 18 false diagnoses and 285 proper diagnoses. The artificial neural network resulted in 15 false diagnoses and 288 proper diagnoses. The support vector machines resulted in 11 false diagnoses and 292 proper diagnoses. Finally, the k-nearest neighbors resulted in 32 false diagnoses and 271 proper diagnoses. Figure 5 shows the variables distribution in univariate or multivariate forms, with a chart illustrating each input variable's issuance. All variables are dispersed in an orderly manner, where the Univariate variable includes the analysis of only one variable. At the same time, the multivariate provides for the study of two or more variables according to the input dataset. Figure 6 illustrates the working mechanism in which heart disease data is analyzed. Table 1 demonstrates the effects obtained by utilizing the five most utilized techniques on the same dataset to measure the performance of each technique. However, this table illustrates that the most useful performance technique is support vector machines, which achieved an accuracy of 96%. In contrast, the most inadequate performance is the k-nearest neighbors technique, which earned an accuracy of 89%. Moreover, artificial neural network random forest, and logistic regression techniques have good and acceptable performance in diagnosing heart disease. Figure 7 illustrates the effects of the techniques involved in the diagnosis process. Also, Table 1 illustrates the most useful technique in terms of performance. The most effective technique for diagnosing Cleveland Clinic dataset has been identified in same table. Besides, Table 2 compares this work with previous studies that used machine learning techniques to analyze the same dataset utilized. In this scenario, four mathematical equations are relied upon to find the effects of the techniques, which are: Accuracy: indicates the overall performance of the techniques (1), Sensitivity is the ratio of heart patients to the total number of patients in the dataset and it gives a positive confirmation of the patient's diagnosis (2), Specificity: is the diagnosis of healthy people as unhealthy (3), and Precision is the marker of a positive diagnosis made by the techniques (4). The last metric is the F1 score, which measures the probability that a positive diagnosis is correct (5).

$$Accuracy = \frac{TP+TN}{TN+TP+FN+FP} \times 100\% \quad (1)$$

$$Sensitivity = \frac{TP}{TN+TP} \times 100\% \quad (2)$$

$$Specificity = \frac{TN}{TN+FP} \times 100\% \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (4)$$

$$F - measure = \frac{2 \times precision \times sensitivity}{pprecision+sensitivity} \quad (5)$$

where TP is true positive, TN is true negative, FN is false negative, and FP is false positive.

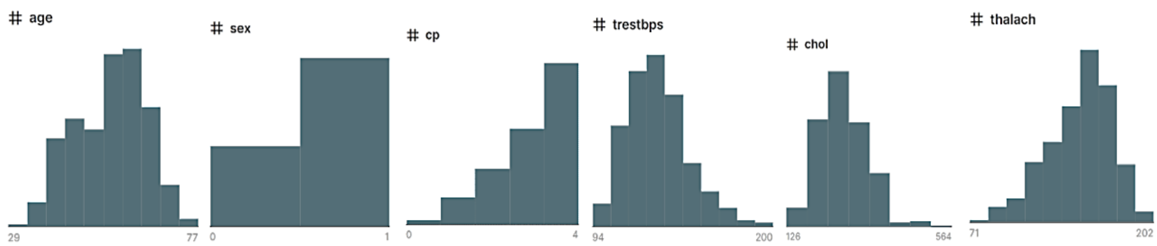


Figure 4. The distribution of variables (age, sex, cp, *trestbps*, *chol*, and *thalach*)

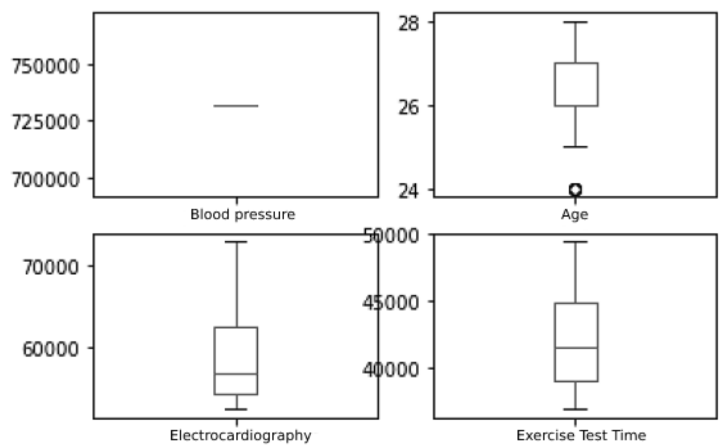


Figure 5. Univariate or multivariate variables (Age, blood pressure, electrocardiography, exercise test time)

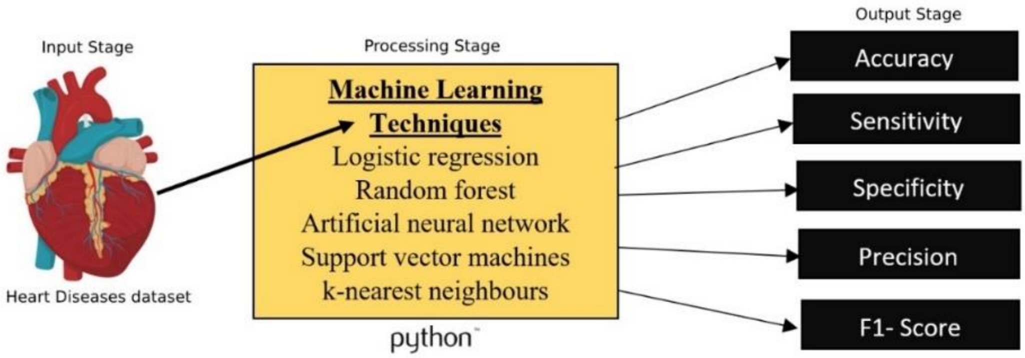


Figure 6. Mechanism of work for diagnosing heart disease dataset

Table 1. Comparison of the effects of the executed techniques

Techniques	Diagnoses Proper	Diagnoses False	Accuracy	Sensitivity	Specificity	Precision	F1-score	Processing time (Second)
Logistic regression	280	23	92%	94%	95%	95%	94%	412
Random forest	285	18	94%	96%	94%	96%	95%	490
Artificial neural network	288	15	95%	95%	94%	98%	95%	501
Support vector machines	292	11	96%	97%	92%	98%	96%	317
k-nearest neighbors	271	32	89%	91%	98%	94%	92%	388

Bold indicates that these effects are the most useful.

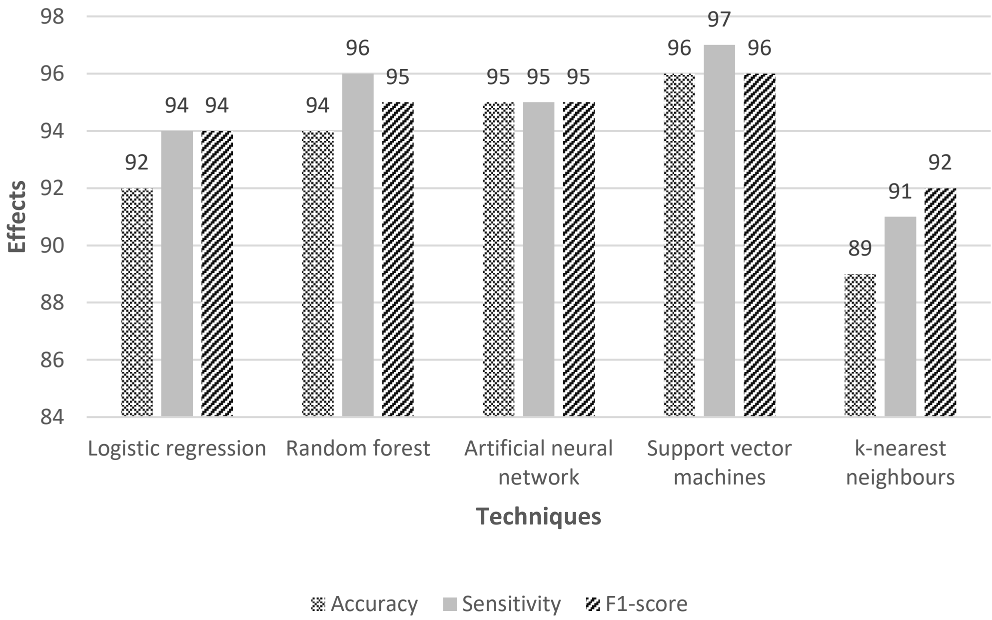


Figure 7. The effects of the techniques applied in the diagnosis of dataset

Table 2. Comparison between the current study and previous studies through the same dataset

Articles	Dataset	Best Techniques	Accuracy
Patel *et al.* [55]	Data Cleveland Clinic dataset set	J48 decision tree	56%
Nassif *et al.* [56]		Naïve Bayes	84%
Terrada *et al.* [57]		Artificial neural network	91%
Akella and Akella [58]		Artificial neural network	93%
This work		Support vector machines	96%

4. CONCLUSION AND FUTURE WORK

The application of machine learning techniques is a common thing for the detection and analysis of diseases due to the predictive ability of these techniques, as it is possible to achieve a prior diagnosis that helps healthcare workers to create decisions that assist save the lives of citizens. The major reason why the authors prefer heart disease is that heart disease is a risk factor in human life, which is considered one of the most dangerous diseases that gives a high mortality rate all over the world. This article applies the five most common and used techniques in analyzing and diagnosing heart diseases. All the used techniques are compared with each other in terms of accuracy, and the most acceptable and worst practice in terms of performance are determined. The artificial neural network technique has taken the most time to implement, reaching 501 seconds, which is considered the most time to implement in this work. While the support vector machines have achieved a short time in the work as it reached 317 seconds, which is the shortest time that is achieved in this work. It is concluded that the artificial neural network technique gives solutions in a long time, and this leads to the difficulty of benefiting from it in the future. Although this technique is difficult for various reasons, but its application is increasing because it gives clear and acceptable results in performance. As for the support vector machines gave the highest accuracy among the applied techniques and in less time; that is, it is considered the fastest applied technique and the most appropriate in diagnosing heart diseases. The summarized literature in this article confirmed that machine learning techniques are necessary and should be applied in the healthcare service because of their high ability to assist in decision-making. In the future, a survey will be conducted on the application of machine learning techniques in the medical field and how they are used. Also, the focus will be on implementing machine learning techniques for analyzing and diagnosing a large medical dataset to create predictions.

REFERENCES

[1] R. Yang, F. R. Yu, P. Si, Z. Yang, and Y. Zhang, "Integrated blockchain and edge computing systems: a survey, some research issues and challenges," *IEEE Communications Surveys and Tutorials*, vol. 21, no. 2, pp. 1508–1532, 2019, doi: 10.1109/COMST.2019.2894727.

[2] H. Valavi, P. J. Ramadge, E. Nestler, and N. Verma, "A 64-Tile 2.4-Mb in-memory-computing CNN accelerator employing charge-domain compute," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 6, pp. 1789–1799, Jun. 2019, doi: 10.1109/JSSC.2019.2899730.

[3] M. Alrazgan, "Internet of medical things and edge computing for improving healthcare in smart cities," *Mathematical Problems in Engineering*, vol. 2022, pp. 1–10, Mar. 2022, doi: 10.1155/2022/5776954.

[4] H. J. de Moura Costa, C. A. da Costa, R. da Rosa Righi, and R. S. Antunes, "Fog computing in health: a systematic literature review," *Health and Technology*, vol. 10, no. 5, pp. 1025–1044, Sep. 2020, doi: 10.1007/s12553-020-00431-8.

[5] L. Zhao, "Event prediction in the big data era: a systematic survey," *ACM Computing Surveys*, vol. 54, no. 5, pp. 1–37, Jun. 2022, doi: 10.1145/3450287.

[6] K. Aggarwal *et al.*, "Has the future started? the current growth of artificial intelligence, machine learning, and deep learning," *Iraqi Journal for Computer Science and Mathematics*, pp. 115–123, Jan. 2022, doi: 10.52866/ijcsm.2022.01.01.013.

[7] R. Vaishya, M. Javaid, I. H. Khan, and A. Haleem, "Artificial intelligence (AI) applications for COVID-19 pandemic," *Diabetes and Metabolic Syndrome: Clinical Research & Reviews*, vol. 14, no. 4, pp. 337–339, Jul. 2020, doi: 10.1016/j.dsx.2020.04.012.

[8] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: a survey," *Heliyon*, vol. 4, no. 11, Nov. 2018, doi: 10.1016/j.heliyon.2018.e00938.

[9] V. Kaul, S. Enslin, and S. A. Gross, "History of artificial intelligence in medicine," *Gastrointestinal Endoscopy*, vol. 92, no. 4, pp. 807–812, Oct. 2020, doi: 10.1016/j.gie.2020.06.040.

[10] Amisha, P. Malik, M. Pathania, and V. Rathaur, "Overview of artificial intelligence in medicine," *Journal of Family Medicine and Primary Care*, vol. 8, no. 7, 2019, doi: 10.4103/jfmpc.jfmpc_440_19.

[11] M. M. Mijwil, "Implementation of machine learning techniques for the classification of lung x-ray images used to detect COVID-19 in humans," *Iraqi Journal of Science*, pp. 2099–2109, Jul. 2021, doi: 10.24996/ijs.2021.62.6.35.

[12] M. Elbadawi, S. Gaisford, and A. W. Basit, "Advanced machine-learning techniques in drug discovery," *Drug Discovery Today*, vol. 26, no. 3, pp. 769–777, Mar. 2021, doi: 10.1016/j.drudis.2020.12.003.

[13] M. M. Mijwil and K. Aggarwal, "A diagnostic testing for people with appendicitis using machine learning techniques," *Multimedia Tools and Applications*, vol. 81, no. 5, pp. 7011–7023, Feb. 2022, doi: 10.1007/s11042-022-11939-8.

[14] F. M. Rammo and M. N. Al-Hamdani, "Detecting the speaker language using CNN deep learning algorithm," *Iraqi Journal for Computer Science and Mathematics*, pp. 43–52, Jan. 2022, doi: 10.52866/ijcsm.2022.01.01.005.

[15] S. Singh, J. Cha, T. Kim, and J. Park, "Machine learning based distributed big data analysis framework for next generation web in IoT," *Computer Science and Information Systems*, vol. 18, no. 2, pp. 597–618, 2021, doi: 10.2298/CSIS200330012S.

[16] H. Kumar, P. J. Soh, and M. A. Ismail, "Big data streaming platforms: a review," *Iraqi Journal for Computer Science and Mathematics*, pp. 95–100, Apr. 2022, doi: 10.52866/ijcsm.2022.02.01.010.

[17] C. Shorten, T. M. Khoshgoftaar, and B. Furht, "Deep learning applications for COVID-19," *Journal of Big Data*, vol. 8, no. 1, Dec. 2021, doi: 10.1186/s40537-020-00392-9.

[18] U. B. Patayon and R. V Crisostomo, "Peanut leaf spot disease identification using pre-trained deep convolutional neural network," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 3, pp. 3005–3012, Jun. 2022, doi: 10.11591/ijece.v12i3.pp3005-3012.

[19] A. K. Faieq and M. M. Mijwil, "Prediction of heart diseases utilising support vector machine and artificial neural network," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 26, no. 1, pp. 374–380, Apr. 2022, doi: 10.11591/ijeecs.v26.i1.pp374-380.

[20] A. Sharma, S. Rani, and D. Gupta, "Artificial intelligence-based classification of chest x-ray images into COVID-19 and other infectious diseases," *International Journal of Biomedical Imaging*, vol. 2020, pp. 1–10, Oct. 2020, doi: 10.1155/2020/8889023.

[21] S. R. Nayak, D. R. Nayak, U. Sinha, V. Arora, and R. B. Pachori, "Application of deep learning techniques for detection of

COVID-19 cases using chest X-ray images: a comprehensive study," *Biomedical Signal Processing and Control*, vol. 64, Feb. 2021, doi: 10.1016/j.bspc.2020.102365.
[22] T. A. Assegie, P. K. Rangarajan, N. K. Kumar, and D. Vigneswari, "An empirical study on machine learning algorithms for heart disease prediction," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 3, pp. 1066–1073, Sep. 2022, doi: 10.11591/ijai.v11.i3.pp1066-1073.
[23] M. Mijwil and B. Salman Shukur, "A scoping review of machine learning techniques and their utilisation in predicting heart diseases," *Ibn AL-Haitham Journal for Pure and Applied Sciences*, vol. 35, no. 3, pp. 175–189, Jul. 2022, doi: 10.30526/35.3.2813.
[24] M. Bourrier, T. W. Ferguson, J. M. Embil, C. Rigatto, P. Komenda, and N. Tangri, "Peripheral artery disease: its adverse consequences with and without CKD," *American Journal of Kidney Diseases*, vol. 75, no. 5, pp. 705–712, May 2020, doi: 10.1053/j.ajkd.2019.08.028.
[25] H. Yang, T. Tian, D. Wu, D. Guo, and J. Lu, "Prevention and treatment effects of edible berries for three deadly diseases: Cardiovascular disease, cancer and diabetes," *Critical Reviews in Food Science and Nutrition*, vol. 59, no. 12, pp. 1903–1912, Jul. 2019, doi: 10.1080/10408398.2018.1432562.
[26] M. M. Fuchs *et al.*, "Cardiovascular surgery in Turner syndrome - early outcome and long-term follow-up," *World Journal of Cardiology*, vol. 12, no. 3, pp. 97–106, Mar. 2020, doi: 10.4330/wjc.v12.i3.97.
[27] M. S. Sabatine *et al.*, "Percutaneous coronary intervention with drug-eluting stents versus coronary artery bypass grafting in left main coronary artery disease: an individual patient data meta-analysis," *The Lancet*, vol. 398, no. 10318, pp. 2247–2257, Dec. 2021, doi: 10.1016/S0140-6736(21)02334-5.
[28] S. Saeed, U. Waje-Andreassen, and P. M. Nilsson, "The association of the metabolic syndrome with target organ damage: focus on the heart, brain, and central arteries," *Expert Review of Cardiovascular Therapy*, vol. 18, no. 9, pp. 601–614, Sep. 2020, doi: 10.1080/14779072.2020.1807327.
[29] M. Kivimäki and A. Steptoe, "Effects of stress on the development and progression of cardiovascular disease," *Nature Reviews Cardiology*, vol. 15, no. 4, pp. 215–229, Apr. 2018, doi: 10.1038/nrcardio.2017.189.
[30] K. Davies, E. Dures, and W.-F. Ng, "Fatigue in inflammatory rheumatic diseases: current knowledge and areas for future research," *Nature Reviews Rheumatology*, vol. 17, no. 11, pp. 651–664, Nov. 2021, doi: 10.1038/s41584-021-00692-1.
[31] A. Alfaddagh *et al.*, "Inflammation and cardiovascular disease: from mechanisms to therapeutics," *American Journal of Preventive Cardiology*, vol. 4, Dec. 2020, doi: 10.1016/j.ajpc.2020.100130.
[32] G. A. Roth, G. A. Mensah, and V. Fuster, "The global burden of cardiovascular diseases and risks," *Journal of the American College of Cardiology*, vol. 76, no. 25, pp. 2980–2981, Dec. 2020, doi: 10.1016/j.jacc.2020.11.021.
[33] S. S. Khan *et al.*, "Association of body mass index with lifetime risk of cardiovascular disease and compression of morbidity," *JAMA Cardiology*, vol. 3, no. 4, p. 280, Apr. 2018, doi: 10.1001/jamacardio.2018.0022.
[34] K. Wójcik-Cichy, E. Koślińska-Berkan, and A. Piekarska, "The influence of NAFLD on the risk of atherosclerosis and cardiovascular diseases," *Clinical and Experimental Hepatology*, vol. 4, no. 1, pp. 1–6, 2018, doi: 10.5114/ceh.2018.73155.
[35] W. D. Strain and P. M. Paldánius, "Diabetes, cardiovascular disease and the microcirculation," *Cardiovascular Diabetology*, vol. 17, no. 1, Apr. 2018, doi: 10.1186/s12933-018-0703-2.
[36] C. Andersson and R. S. Vasan, "Epidemiology of cardiovascular disease in young individuals," *Nature Reviews Cardiology*, vol. 15, no. 4, pp. 230–240, Apr. 2018, doi: 10.1038/nrcardio.2017.154.
[37] M. Nikpouraghdam *et al.*, "Epidemiological characteristics of coronavirus disease 2019 (COVID-19) patients in IRAN: a single center study," *Journal of Clinical Virology*, vol. 127, Jun. 2020, doi: 10.1016/j.jcv.2020.104378.
[38] A. J. Einstein *et al.*, "International impact of COVID-19 on the diagnosis of heart disease," *Journal of the American College of Cardiology*, vol. 77, no. 2, pp. 173–185, Jan. 2021, doi: 10.1016/j.jacc.2020.10.054.
[39] S. Lalmuanawma, J. Hussain, and L. Chhakchhuak, "Applications of machine learning and artificial intelligence for COVID-19 (SARS-CoV-2) pandemic: a review," *Chaos, Solitons and Fractals*, vol. 139, Oct. 2020, doi: 10.1016/j.chaos.2020.110059.
[40] M. M. Ahsan and Z. Siddique, "Machine learning-based heart disease diagnosis: a systematic literature review," *Artificial Intelligence in Medicine*, vol. 128, Jun. 2022, doi: 10.1016/j.artmed.2022.102289.
[41] B. X. Tran *et al.*, "The current research landscape of the application of artificial intelligence in managing cerebrovascular and heart diseases: a bibliometric and content analysis," *International Journal of Environmental Research and Public Health*, vol. 16, no. 15, Jul. 2019, doi: 10.3390/ijerph16152699.
[42] Y. Niu and K. Andrei, "Identification method of power internet attack information based on machine learning," *Iraqi Journal for Computer Science and Mathematics*, pp. 1–7, Feb. 2022, doi: 10.52866/ijcsm.2022.02.01.001.
[43] T. Leiner *et al.*, "Machine learning in cardiovascular magnetic resonance: basic concepts and applications," *Journal of Cardiovascular Magnetic Resonance*, vol. 21, no. 1, Dec. 2019, doi: 10.1186/s12968-019-0575-y.
[44] F. W. Damen, D. T. Newton, G. Lin, and C. J. Goergen, "Machine learning driven contouring of high-frequency four-dimensional cardiac ultrasound data," *Applied Sciences*, vol. 11, no. 4, Feb. 2021, doi: 10.3390/app11041690.
[45] Z. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei, and A. A. Yarifard, "Computer aided decision making for heart disease detection using hybrid neural network-genetic algorithm," *Computer Methods and Programs in Biomedicine*, vol. 141, pp. 19–26, Apr. 2017, doi: 10.1016/j.cmpb.2017.01.004.
[46] AHA, "What is Cardiovascular Disease?, heart attack and stroke symptoms," *American Heart Association*, 2017. https://www.heart.org/en/health-topics/consumer-healthcare/what-is-cardiovascular-disease (accessed May 1, 2022).
[47] M. J. Ankenbrand, L. Shainberg, M. Hock, D. Lohr, and L. M. Schreiber, "Sensitivity analysis for interpretation of machine learning based segmentation models in cardiac MRI," *BMC Medical Imaging*, vol. 21, no. 1, Dec. 2021, doi: 10.1186/s12880-021-00551-1.
[48] Y. Pan, M. Fu, B. Cheng, X. Tao, and J. Guo, "Enhanced deep learning assisted convolutional neural network for heart disease prediction on the internet of medical things platform," *IEEE Access*, vol. 8, pp. 189503–189512, 2020, doi: 10.1109/ACCESS.2020.3026214.
[49] C. Ricciardi *et al.*, "Assessing cardiovascular risks from a mid-thigh CT image: a tree-based machine learning approach using radiodensitometric distributions," *Scientific Reports*, vol. 10, no. 1, Dec. 2020, doi: 10.1038/s41598-020-59873-9.
[50] A. Helwan and D. Uzun Ozsahin, "Sliding window based machine learning system for the left ventricle localization in MR cardiac images," *Applied Computational Intelligence and Soft Computing*, vol. 2017, pp. 1–9, 2017, doi: 10.1155/2017/3048181.
[51] F. I. Alarsan and M. Younes, "Analysis and classification of heart diseases using heartbeat features and machine learning algorithms," *Journal of Big Data*, vol. 6, no. 1, Dec. 2019, doi: 10.1186/s40537-019-0244-x.
[52] M. M. Ali, B. K. Paul, K. Ahmed, F. M. Bui, J. M. W. Quinn, and M. A. Moni, "Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison," *Computers in Biology and Medicine*, vol. 136, Sep. 2021,

doi: 10.1016/j.compbiomed.2021.104672.
[53] J. H. Joloudari *et al.*, "FCM-DNN: diagnosing coronary artery disease by deep accuracy fuzzy C-means clustering model," *Mathematical Biosciences and Engineering*, vol. 19, no. 4, pp. 3609–3635, 2022, doi: 10.3934/mbe.2022167.
[54] A. Cook, "Cleveland clinic foundation heart disease," *Kaggle, 2020*. Accessed May 1, 2022. [Online]. Available: https://www.kaggle.com/datasets/alexisbcook/cleveland-clinic-foundation-heart-disease
[55] Jaymin Patel, Tejal Upadhyay, and Samir B Patel, "Heart disease prediction using machine learning and data mining technique," *International Journal of Computer Science and Communication*, vol. 7, no. 1, pp. 129–137, 2016.
[56] A. B. Nassif, O. Mahdi, Q. Nasir, M. A. Talib, and M. Azzeh, "Machine learning classifications of coronary artery disease," in *2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (ISAI-NLP)*, Nov. 2018, pp. 1–6, doi: 10.1109/iSAI-NLP.2018.8692942.
[57] O. Terrada, S. Hamida, B. Cherradi, A. Raihani, and O. Bouattane, "Supervised machine learning based medical diagnosis support system for prediction of patients with heart disease," *Advances in Science, Technology and Engineering Systems Journal*, vol. 5, no. 5, pp. 269–277, 2020, doi: 10.25046/aj050533.
[58] A. Akella and S. Akella, "Machine learning algorithms for predicting coronary artery disease: efforts toward an open source solution," *Future Science OA*, vol. 7, no. 6, Jul. 2021, doi: 10.2144/fsoa-2020-0206.

BIOGRAPHIES OF AUTHORS

Ban Salman Shukur received the B.Eng. degree in computer science from university of Technology, Baghdad, Iraq in 1987. M.Sc. in Information Technology from UUM, Malaysia 2012 and Ph.D. degrees in Information Technology 2021 from UTeM, Malaysia. Working as a senior lecturer in Computer Science Department at Baghdad college of Economic Sciences University, Baghdad, Iraq. Experience in teaching Undergraduate Computer Science students since 2012 till now. Member of editorial board in Internet of Things and Cloud Computing (IOTCC), Science publication group, New York, United States since November 2021. She can be contacted at email: Dr_bansalman@baghdadcollege.edu.iq.

Maad M. Mijwil is an Iraqi Academician; he is born in Baghdad, Iraq, in 1987. He received his B.Sc. degree in software engineering from Baghdad college of economic sciences university, Iraq, in 2009. He received a M.Sc. degree in 2015 from the Computer Science Department, University of Baghdad in the field of wireless sensor networks, Iraq. Currently, he is working as a Lecturer and an academic member of staff in the Computer Techniques Engineering Department at Baghdad College of Economic Sciences University, Iraq. He has over ten years of experience in teaching and guiding projects for undergraduates. He has authored more than 80 publications, including papers/ chapters (published 51 peer reviewed papers in national/international conferences and journals), preprints, presentations, and posters. He is also an editor in more than 10 international/national journals and a reviewer in more than 80 international/national journals. He has served on technical program committees for many prestigious conferences. Also, he graduated from Publons academy as a peer reviewer. His Google citations are over 400. His research interests include artificial intelligence, machine learning, deep learning, wireless sensor networks, genetic algorithm. He can be contacted at email: mr.maad.alnaimiy@baghdadcollege.edu.iq.