

A Deep Learning Technique in Cardio Vascular Disease Prediction Using Convolutional Neural Network

Asma'u Muhammad Sani¹, Rabi Mustapha² and Khalid Haruna³

^{1,2,3}Department of Computer Science, Kaduna State University, Kaduna, Nigeria
asmausn@gmail.com¹, rabichubu@kasu.edu.ng², khalidharuna@kasu.edu.ng³

Abstract

Heart disease or Cardio-vascular disease is considered as a fatal illness that continues to increase faster in our modern world, in 2012, around 17.5 million individuals kicked the bucket from coronary illness, implying that it comprises of the 31% of every single worldwide passing. Besides, coronary illness loss of life rises each year. This research aims at applying and improving a deep learning technique to detect Cardio-Vascular Disease using 1D Convolutional Neural Network, to also evaluate the performance of the proposed system and compare it with the existing system. Since medical big data has been increasing daily and data storage costs decreasing, deep learning algorithms can play an important role in processing these medical data and predicting diseases. With the help of the deep learning (CNN) Classifier algorithm, the research was able to build a deep-learning model. We obtained a precision of 83.9% a recall of 83.9% and an F-Measure of 83.9% which shows that our proposed system outperforms the existing system.

Keywords: cardiovascular disease, deep learning, convolutional neural network, Framingham dataset

1. Introduction

Heart disease or Cardio-vascular disease is considered as a fatal illness that continues to increase faster in our modern world. World Health Organization (WHO) expects that the number of deaths will add up by 24.5 in 2030 (Nabaouia, Meryem & Bouabid 2019). Many factors may cause heart disease. The changing in lifestyle, smoking, food habits, physical activity, obesity, and diabetes also biochemical factors like blood pressure or glycemia.

In 2012, around 17.5 million individuals kicked the bucket from coronary illness, implying that it comprises of the 31% of every single worldwide passing. Besides, coronary illness loss of life rises each year. It is relied upon to develop more than 23.6 million by 2030 Rubini et al., (2022). The exploration from the January 2017 demonstrated that the main source of death worldwide is cardiovascular infections. The cardiovascular malady is considered as a world's biggest killer and is currently taking the top position in the record of ten reasons for passing in the previous 15 years and in 2015 was numeration for fifteen million passing. Various human lives could be spared by diagnosing on schedule (Hossam, 2019). Along these lines, diagnosing the syndrome is significant and an exceptionally muddled undertaking. For this reason, it's compulsory to record critical heart carriage depending on each type of heart disease and provide a system that helps doctors making correct and efficient decisions while diagnosing. Medical diagnosis is, in fact, a classification mission where a doctor tries identifying the defect by analyzing the values of a number of characteristics. That task is generally based on a variety of algorithms.

On one hand, traditional ways on diagnosing heart disease are ECG, occultation and taking measures of blood pressure, blood sugar, and cholesterol. However, those technics are expensive and consuming in time. On the other hand, Machine learning algorithms (Support Vector Machine, Bayes Naïve, K-Nearest Neighbour etc) makes the diagnosis of CVD an opportune solution that reduces extremely the time processing and increases the prediction performance. Previous works have used machine learning algorithms to diagnosis heart disease. Unfortunately, the accuracy, sensitivity, and specificity were so poor. Not only that, but the performance metrics that they have used to evaluate their results insufficient. The remaining part of this paper is organized as follows; Section 2 contains related works, Section 3 presents the methodology used in the research while Section 4 presents the results and discussion obtained and finally Section 5 is the conclusion and future work

2. Related Work

The traditional healthcare system uses expert's knowledge to diagnose and predict medical issues. However, the experts can miss data new patterns and the definition of feature space. As a solution, machine and deep learning approaches can predict and extract those patterns and feature space automatically (Yazhini, & Loganathan, 2019). The increasing of digital healthcare data has driven the data medical research based on machine and deep learning techniques. Those techniques are powerful for big data and perform well in feature representation and pattern recognition Yu et al., (2019). Dataset is used to feed a neural network, which delivers a prediction, a diagnosis, a treatment or a follow up in the healthcare system. For example, deep learning approaches have been used in applications related to visual feature extraction, speech identification, and textual analysis. Healthcare monitoring involves the use of wearable devices and smartphones, which transfer data to mobile apps or a medical center. That data can be complex and noisy, where deep learning model offers a novel and efficient solutions for healthcare data. Thus, the goal of using deep learning is representation, where input data features can be represented by training multilayer neural network effectively Yu et al., (2019).

Deep learning builds neural networks based on the structure of the brain by being inspired from neuroscience. Deep neural networks (DNN) approaches can learn classification and feature representation from raw data Hung et al., (2019). A deep architecture is accomplished by adding more hidden layers to the neural network, which allows capturing nonlinear relationships Ravi et al., (2017). Traditional deep learning applications have been implemented using only one modality (i.e., or multimedia data) such as text, image, audio or video. In multimodal applications, deep learning solutions analyze all the information from different modalities to find logical connections between all of them to make decisions that are more accurate (Chen, 2019). Deep learning approaches have influenced medical applications.

They are useful to handle multi-modality data, learn relationship among data, extract predictive models, and analyze data patterns in real time. For example, deep learning can capture patterns of patients from clinical data to make predictions or to diagnose diseases. By monitoring a patient for a period, accumulated clinical data can be analyzed by specialized personnel to make prognosis. DNN allow identification of clinical notes, where it can extract information from unstructured text. In addition, early warning can be done since DNN will help to detect and diagnose a disease early as well as to do preventive treatment, consequently reducing healthcare cost and saving lives Yu, et al., (2019). Deep learning combined with large volume of data and GPU enable to explore medical data to achieve precision medicine Yu, et al., (2019). In medical image, it is hard to extract handcrafted features, whereas deep learning simplifies that extraction and improves the precision in diagnosis. Used deep learning structures in

healthcare are CNN, recurrent neural networks (RNNs), and Auto Encoders (AEs) (Yazhini, & Loganathan, 2019). CNNs are based on their shared-weights convolutional kernels as well as characteristics of translation invariance Li, et al.(2019). This is the deep learning model more used in clinical applications (e.g., image diagnosis, electrocardiogram monitoring, image segmentation, and diagnose from clinical notes Li, et al.,(2019). It detects anomalous findings in different modality images such as histology or tomography scans in large-scale images from thousands of patients. In addition, low-level features are learned and fine-tuned using different images and networks (Lu & Harrison, 2018). Using CNN has helped to reduce the vanishing gradient problem, fortify feature propagation, allow feature reuse, and reduce the number of parameters (Huang, Liu, Maaten & Weinberger,2017). RNNs, in turn, are a powerful tool for varying length sequences such as audios, videos, and sentences (Guo, Wang, Wang, 2019).

In the research conducted by Dritsas, (Alexiou, & Moustakas, 2022) the researchers evaluated risk for CVD occurrence on participants (men and women) especially aged older than 50 years with the aid of Machine Learning (ML) models; Logistic Regression classifier, Naive Bayes, Support Vector Machine (SVM) and Random Forest they obtained 72.1% Accuracy, recall and 78.4% AUC. The limitation of the work is the utilization of deep learning models and techniques as they can provide complex decision boundaries and thus better fit the training data. Rubini, P. et al., (2022) conducted a Comparative analysis of machine learning techniques like Random Forest (RF), Logistic Regression, Support Vector Machine (SVM), and Naïve Bayes in the classification of cardiovascular disease and obtained the following results Random Forest 84.81%, SVM 74.05%, Naïve Bayes 54.08%. The future direction of this work is to use similar prediction systems can be built in calculating correlation between heart diseases and other diseases. Also new algorithms can be used to achieve increased accuracy. Amzad, et al., (2021) also conducted a comparative assessment of three machine learning techniques namely Random Forest (RF), Decision Tree (DT), and Logistic Regression (LR) and obtained an accuracy score of 92.10%. A web application that integrates these methods and uses a larger dataset than the one used in this study might be developed in the future. Aqsa, et al., (2021) proposed a **Machine Learning based Cardiovascular Disease Diagnosis (MaLCaDD)** framework for the effective prediction of cardiovascular diseases with high precision with the following methodology Ensemble of Logistic Regression and K-Nearest Neighbor (KNN) and obtained an accuracy of 99.1%, 98.0% and 95.5 % are achieved respectively for the two classifiers. For the future work they intended to share the real time evaluation results of MaLCaDD on different patients. (Sachin, Ahmad, Ishan, Bina, & Xi, 2020) proposed a hypothesized that machine learning (ML) could be used for gut microbiome-based diagnostic screening of CVD using the following classifiers Decision tree, random forest; and SVM and achieved a promising AUC of ≈ 0.70 with a single parameter of stool gut microbiome data. In future, it will be of interest to further calibrate and improve predictive capability of ML modeling by including more samples from different sources or stratifying specific types of CVD incorporated with combinatorial features such as health records. (Hana, & Manal, 2021) proposed an automatic detector of cardiovascular disease using two datasets through a deep learning network and a variety of machine learning classification models as Support Vector Machine (SVM), Decision Tree Classifier (DT), Logistic Regression, K Nearest Neighbor (KNN), and Naïve Bayes (NB) with the following results Random Forest 94%, Gradient Boosting 73% accuracy, 73% Recall, 73% F1-score, and 74% Precision. Varun et al. (2023) proposed an integrated system that uses deep neural network and CBR for detecting the severity of heart disease their proposed system diagnosed cardiovascular disease among patients based on their clinical data using a deep neural network and obtained a detection accuracy of 96.2%. (Khan & Elfouly, 2023) proposed a hybrid one-dimensional convolutional neural network (1D CNN), which uses a large dataset accumulated

from online survey data and selected features using feature selection algorithms, the non-coronary heart disease (no-CHD) and CHD validation data showed an accuracy of 80.1% and 76.9%, respectively

3. Methodology

This paper employed the use of deep learning using convolutional neural network algorithm and the experiment was carried out in WEKA simulation environment.

3.1 Proposed System

The architecture for the proposed system is presented below in figure 1 with five components which includes: Dataset, Data preparation and the classifier.

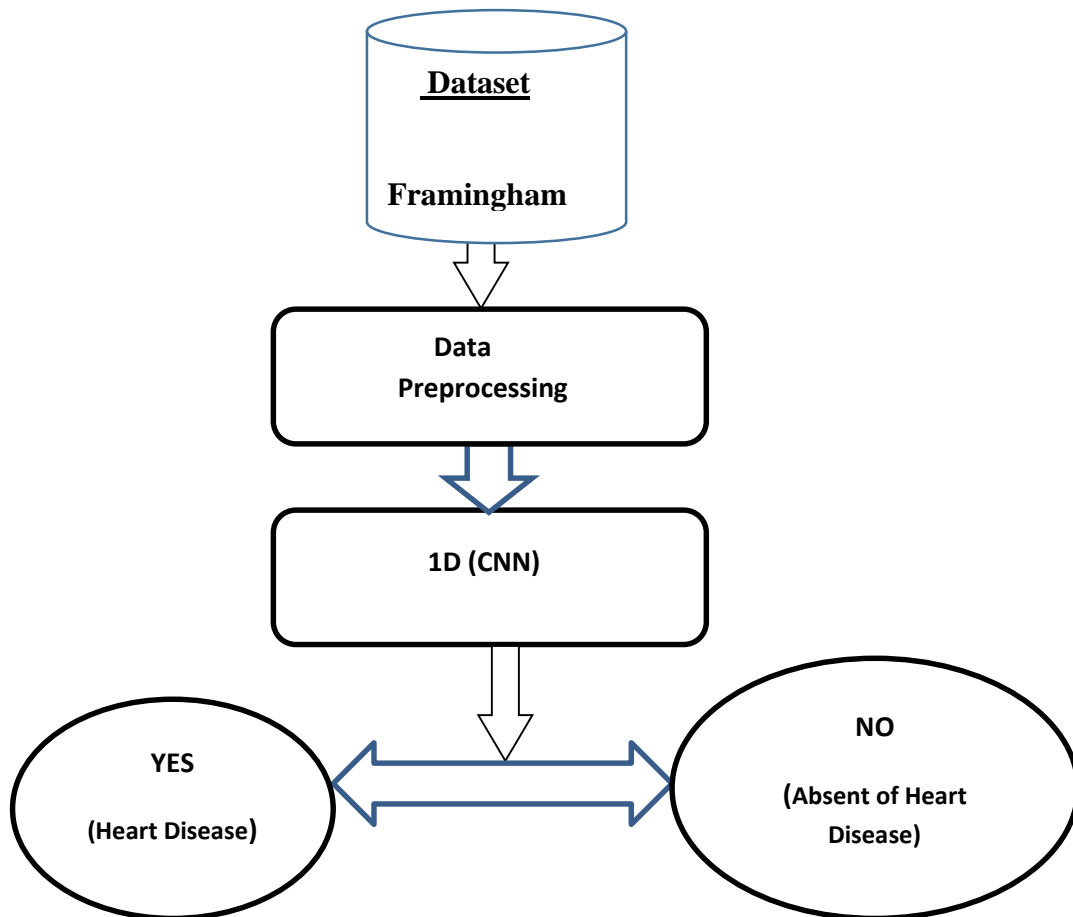


Figure 1 architecture of the proposed system

3.1.1 Dataset

At the course of this thesis, the Framingham dataset will be used as it is one of the frequently used datasets and it contains all the necessary attributes which can help us in the prediction of CVD diseases. The dataset is a rather small subset of possible FHS dataset, having 4240 observations and 16 features. The features are as follows:

- i. sex : the gender of the observations. The variable is a binary named “male” in the dataset.
- ii. age : Age at the time of medical examination in years.
- iii. education : A categorical variable of the participants education, with the levels: Some high school (1), high school/GED (2), some college/vocational school (3), college (4)
- iv. currentSmoker: Current cigarette smoking at the time of examinations
- v. cigsPerDay: Number of cigarettes smoked each day
- vi. BPmeds: Use of Anti-hypertensive medication at exam
- vii. prevalentStroke: Prevalent Stroke (0 = free of disease)
- viii. prevalentHyp: Prevalent Hypertensive. Subject was defined as hypertensive if treated
- ix. diabetes: Diabetic according to criteria of first exam treated
- x. totChol: Total cholesterol (mg/dL)
- xi. sysBP: Systolic Blood Pressure (mmHg)
- xii. diaBP: Diastolic blood pressure (mmHg)
- xiii. BMI: Body Mass Index, weight (kg)/height (m)²
- xiv. heartRate: Heart rate (beats/minute)
- xv. glucose: Blood glucose level (mg/dL)
- xvi. And finally the response variable : + TenYearCHD : The 10 year risk of coronary heart disease(CHD).

3.1.2 Data Pre-processing

A preprocessing step will be carried out in order to clean and extract more useful information from the dataset. The age will be dropped and three new columns will be added representing an age range: young, adult and elder. In the same way, the feature resting BP will be converted into three new columns for lowBP, mediumBP and highBP. Lastly, the Cholesterol feature will also be converted into three categorical columns which determine the risk: low, medium, high. One hot encoding technique will be applied on three features: ChestPainType, RestingECG and ST Slope. At the end of this process, we have a dataset made up of 24 features. A k-fold cross-validation will be carried out with 10 folds for all of the experiments in order to avoid randomness.

3.1.3 Data Classification for Convolution Neural Network

Convolutional neural networks can take two-dimensional data, such as images, as input and are capable of extracting complex features from that data. This information is extracted using filters (kernels). In the training process, the filter weights will be adjusted to carry out an accurate feature map of each class. In the

architecture of a CNN, each convolutional layer must be followed by a pooling layer. The pooling layers are responsible for reducing the overfitting and the number of network parameters so that the computation is not so heavy. The most widely used type of pooling is Max-pooling, which acts by selecting the maximum value of each window. The last layers of the convolutional network have to be dense layers to be able to carry out the classification of the features extracted by the kernels.

The research is going to use the RMSprop optimizer because of the following advantages to medical dataset: RMSprop (Root Mean Square Propagation) offers several advantages that is beneficial when working with a heart disease dataset:

- i. Adaptive Learning Rates: RMSprop adapts the learning rates for each parameter individually based on the magnitude of the gradients. This can be advantageous when dealing with features in a heart disease dataset that may vary significantly in scale.
- ii. Effective Handling of Sparse Gradients: RMSprop can handle sparse gradients well, which may be beneficial if your heart disease dataset has missing or sparse data.
- iii. Robustness to Noisy Gradients: RMSprop's use of a moving average of squared gradients can provide some robustness to noisy or erratic gradients, which may be present in medical datasets.
- iv. Convergence Speed: RMSprop often converges faster than traditional stochastic gradient descent (SGD) because of its adaptive learning rate mechanism. This can be advantageous when dealing with large and complex heart disease datasets.
- v. Less Sensitive to Learning Rate Hyperparameter: RMSprop is less sensitive to the choice of the initial learning rate compared to vanilla SGD. This property can simplify hyperparameter tuning.

When working with a heart disease dataset, experimenting with different optimizers, including RMSprop, and tuning hyperparameters is crucial to find the best-performing model for your specific task.

4. Results and Discussion

These sections of the paper described the analysis of the learned representations in predicting cardiovascular disease using deep learning (1D CNN) and compare the performance of the classifiers with the existing system. At the course of this research, the Framingham dataset was used as it is one of the frequently used datasets and it contains all the necessary attributes which can help us in the prediction of CVD diseases.

Database: Number of instances:

Instances: 367

Attributes: 12

The column number 12 contains the output class which can be 1 (Presence of heart disease) or 0 (Absence of Heart disease).

The experiment was carried out on WEKA 3.9.6 an open source machine learning scripting software. The figure 2 below shows the Home Page of the software with different features like Explorer, Experimenter, KnowledgeFlow, Workbench and Simple CLI.

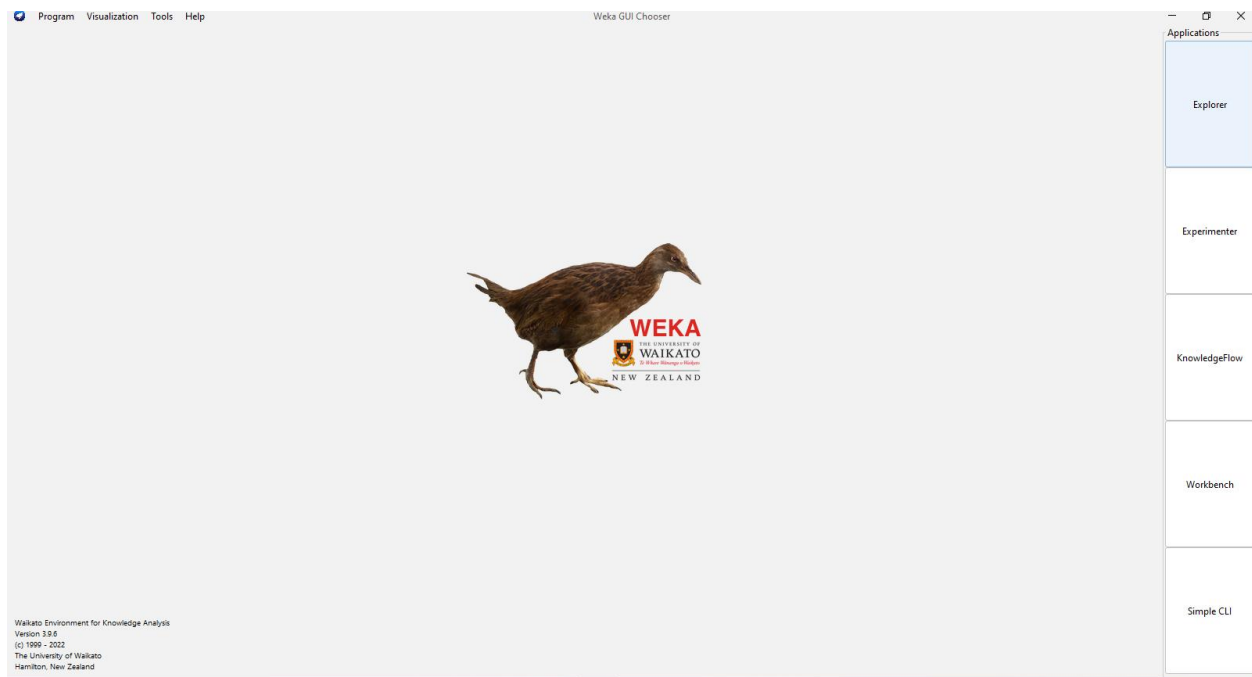


Figure 2: Home page of WEKA 3.9.6 software

The figure below shows the output of the CNN classifier with 84% correctly classified instances and 16% incorrectly classified instances

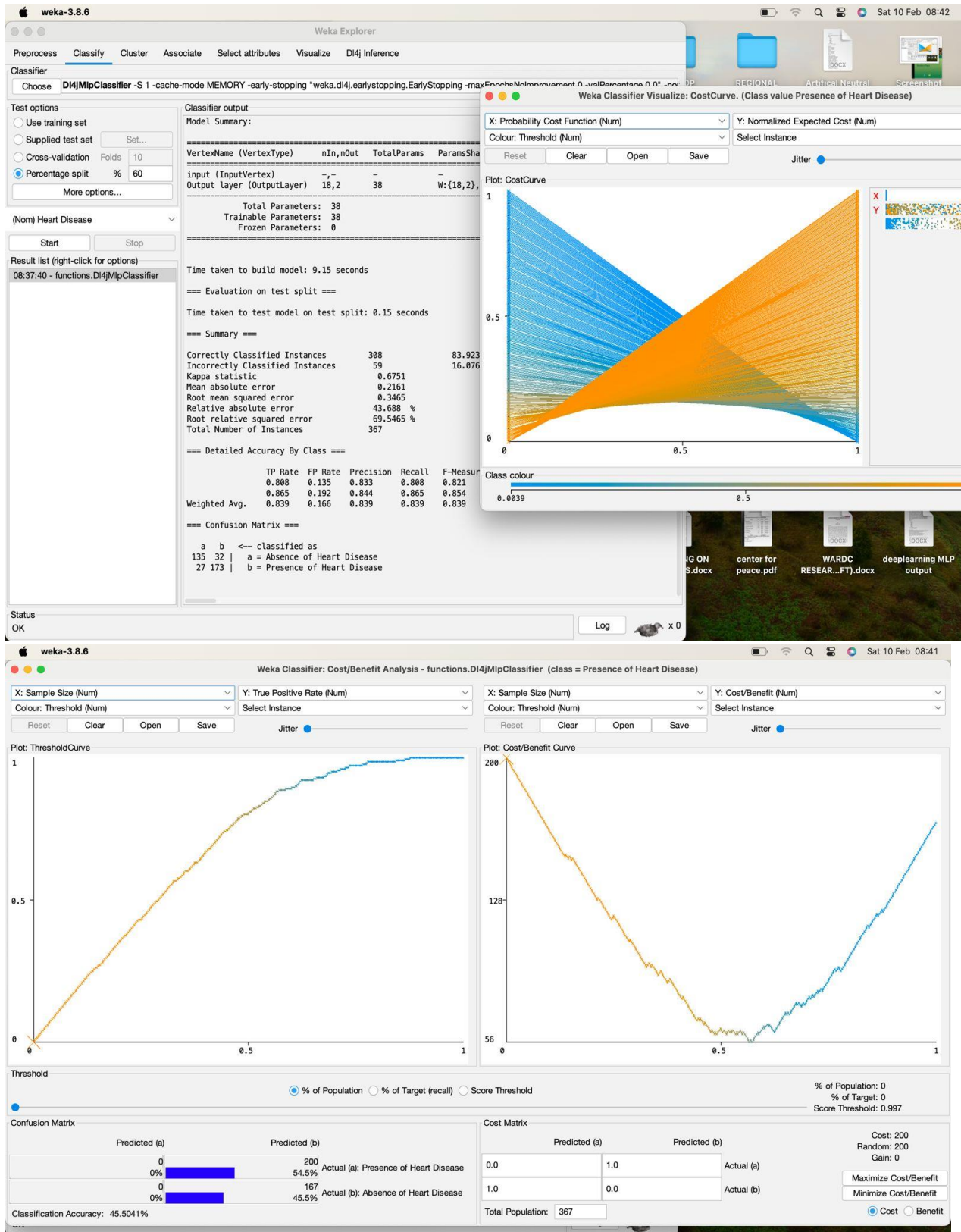


Figure 3 Output results of CNN

From the figures or results above, the two Convolutional Neural Network algorithms performed a classification task and it shows that the CNN algorithm has the highest precision in classifying the patients with cardiovascular heart disease in the class label in the experiment.

Also, the results show that CNN algorithm has the highest detection accuracy. Finally, CNN classifier outperforms the other classifier in carrying out sensitivity, specificity, negative predictive value, false positive rate, and false negative rate.

A Cardiovascular Disease is evaluated by the measure of accuracy, detection rate and F-measure. Precision is the percentage of the total number of patients with coronary artery disease that are properly detected. It is measure with the equation below:

(i). **Accuracy** (Acc) is one of the most commonly used measures for the classification performance, and it is defined as a ratio between the correctly classified samples to the total number of samples and is given by:

$$\text{Accuracy} = \frac{TP}{TP+FP} \dots\dots\dots(1)$$

(ii). Detection Rate or Recall is described as the number of attacks detected by the proposed technique to the total number of attacks truly there (Modi & Jain, 2016).

$$\text{Detection Rate (Recall)} = \frac{TP}{TP+FN} \dots\dots\dots(2)$$

$$\text{(iii). F-Measure} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \dots\dots\dots(3)$$

True Positive (TP): it is the number of patients with coronary artery disease that were correctly classified.

False Positive (FP): this is the number of patients with coronary artery disease that were incorrectly classified as normal.

False Negative (FN): this is the number patients with coronary artery disease that were incorrectly classified.

Table 1 Percentage of Weighted Average of the CNN Classifiers (Proposed system)

Classifier	Precision	Recall	F-Measure	Accuracy
CNN (%)	83.38	83.9	83.9	0.84

Table 1 described the percentages of the weighted average of the convolutional neural network classifiers that were used to perform the experiment.

In this thesis, the existing system developed by (Khan & Elfouly, 2023) was implemented and the outputs are given below:

Table 4.3: Comparison of the Proposed System against the Existing System

	Precision	Recall	F-Measure	Accuracy
Proposed system	83.38	83.9	83.9	0.83
(Khan & Elfouly, 2023)	75.74	78.59	72.68	0.76

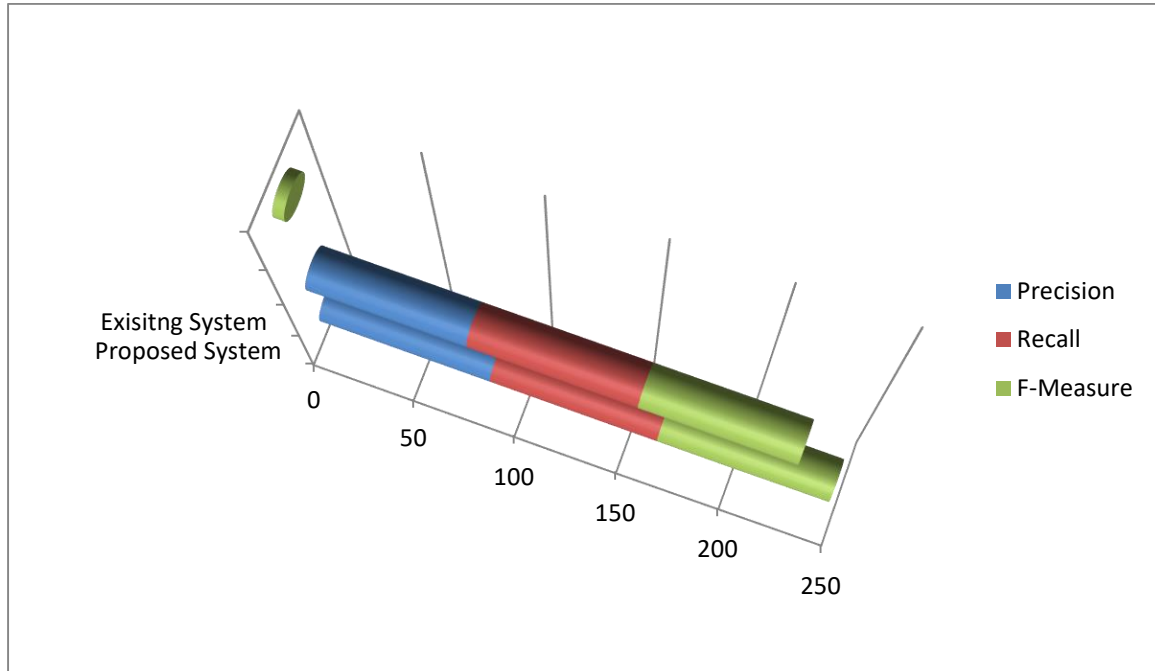


Figure 4 Graph representing the proposed algorithms and existing system

With the above comparisons of both the existing and the proposed system, it is clear that the deep learning model of the proposed algorithms outperforms that of the existing system. It will be observed that precision, recall and the F-measure of the proposed algorithms has more percentage compared to that of the existing system. Figure 4.4 shows a chart that shows the classifiers of both the proposed algorithms (1D CNN) and the existing system, the figure shows how the performance of the proposed algorithms outperforms the existing system.

5. Conclusion

Cardiovascular disease is a life-threatening disease affecting millions of people around the world every year Asadi et al., (2021). Hence, early prediction of cardiovascular disease can benefit patients and healthcare professionals by providing the information they need to minimize death and reduce costs. Since medical big data has been increasing daily and data storage costs decreasing, deep learning algorithms can play an important role in processing these medical data and predicting diseases. With the help of the deep learning (CNN) Classifier algorithm, the research was able to build a deep-learning model. The model was trained and tested by a dataset.

The research can be concluded that deep learning can play an important role in our healthcare system. Traditionally, diagnosis of the disease was performed by standard procedures and doctor's intuitions which had limitations and led to costly expenses, but with machine learning models, diagnosis can be done on large datasets cost-effectively, we obtained a precision of 83.3% a recall of 80.8% and an F-Measure of 82.1% which shows that our proposed system outperforms the existing system.

. In the future this research can be improved by using other unsupervised machine learning (deep learning) like LSTM, using Autoencoder for the feature extraction.

References

- Aditya, M. Prince, K. Himanshu, Arya, & Pankaj, K. (2015). EarlyHeart Disease Prediction Using Data Mining Techniques”, CCSEIT, DMDB, ICBB, MoWiN, AIAP – 2014, pp. 53–59, DOI: 10.5121/csit.2014.4807
- Animesh, H. Subrata, K. Mandal, A. Gupta, Arkomita Mukherjee & Asmita Mukherjee (2017). Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review”, *Advances in Computational Sciences and Technology*, ISSN 0973-6107 Volume 10, Number 7 (2017) pp. 2137-2159
- Asadi, S., Roshan, S. E., & Kattan, M. W. (2021). Random forest swarm optimization-based for heart diseases diagnosis. *Journal of Biomedical Informatics*, 115, 103690. <https://doi.org/10.1016/j.jbi.2021.103690>
- Amzad, M. D. Tahia, T. Sumiaya, K. Evan, A. Hossain, A. Mohammad, M. & Abdulmajeed, A. (2021). Supervised Machine Learning-Based Cardiovascular Disease Analysis and Prediction. *Hindawi Mathematical Problems in Engineering* Volume 2021, Article ID 1792201, 10 pages <https://doi.org/10.1155/2021/1792201>
- Atharv, N. Sanket, B. Aditya, M. & Shamla, M. (2020). Cardiovascular Disease Prediction Using Machine Learning Models. *2020 IEEE Pune Section International Conference (PuneCon) Vishwakarma Institute of Technology, Pune India. Dec 16-18, 2020*
- Aqsa, R. Yawar, R. Farooque, A. Muhammad, Waseem, A. Muhammad, Abdul, R. & AbdulWahab, M. (2021). An Integrated Machine Learning Framework for Effective Prediction of Cardiovascular Disease. *IEEE Access*. Digital Object Identifier 10.1109/ACCESS.2021.3098688
- Bambrick, N. (2022). Support Vector Machines: A simple explanation. KDnuggets. Retrieved November 3, 2022, from <https://www.kdnuggets.com/2016/07/support-vector-machines-simple-explanation.html>
- Bhardwaj, R., Nambiar, A. R., & Dutta, D. (2017). A study of machine learning in Healthcare. 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC). <https://doi.org/10.1109/compsac.2017.164>
- Bharti, R. Khamparia, A. Shabaz, M. Dhiman, G. Pande, S. & Singh, P. (2021). Prediction of heart disease using a combination of machine learning and deep learning,” *Computational Intelligence and Neuroscience*, vol. 2021, Article ID 8387680, 11 pages, 2021.
- Camm, J. D., Cochran, J. J., Fry, M. J., & Ohlmann, J. W. (2021). *Business analytics: Descriptive, predictive, prescriptive*. Cengage.
- Centers for Disease Control and Prevention. (2022, September 8). Heart disease and stroke. Centers for Disease Control and Prevention. Retrieved September 28, 2022, from <https://www.cdc.gov/chronicdisease/resources/publications/factsheets/heart-disease-stroke.htm>
- Chauhan, N. S. (2022, April). Naïve Bayes Algorithm: Everything you need to know. KDnuggets. Retrieved October 29, 2022, from https://www.kdnuggets.com/2020/06/naive-bayes-algorithm-everything.html?hss_channel=tw-1318985240

- Chen, L. (2019). Support Vector Machine — simply explained - towards data science. Support Vector Machine — Simply Explained. Retrieved November 4, 2022, from <https://towardsdatascience.com/support-vector-machine-simply-explained-fee28eba5496>
- Chen, S.-C.: Multimedia deep learning. *IEEE Multimed* **26**(1), 5–7 (2019)
- Chandrasekhar, T. Harsha, S.S.V Pavan, & Vinay, B. (2019). Detection of Heart Diseases Using Data Mining Techniques”, *International Journal of Engineering Sciences & Research Technology*, *IJESRT*: 8(3), March, 2019 ISSN: 2277-9655
- Chaitrali, S. Dangare, S. & Apte, S. (2018). Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques”, *International Journal of Computer Applications* (0975 – 888), Volume 47–No.10, June 2018
- Dinesh, K. Santhosh, K. Arumugaraj, K. & Mareeswari, V. (2018). Prediction of Cardiovascular Disease Using Machine Learning Algorithms. *Proceeding of 2018 IEEE International Conference on Current Trends toward Converging Technologies, Coimbatore, India*
- Dritsas, E., Alexiou, S. and Moustakas, K. (2022). Cardiovascular Disease Risk Prediction with Supervised Machine Learning Techniques. DOI: 10.5220/0011088300003188 In *Proceedings of the 8th International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE 2022)*, pages 315-321. ISBN: 978-989-758-566-1; ISSN: 2184-4984
- Fatima, M. & Pasha, M. (2017). Survey of machine learning algorithms for disease diagnostic. *J Intell Learn Syst Appl*. 2017;9:1–16. <https://doi.org/10.4236/jilsa.2017.91001>
- Foresee Medical (2020) Benefits of Machine Learning in healthcare <https://www.foreseemed.com/blog/machine-learning-in-healthcare>
- Global Atlas on Cardiovascular Disease Prevention and Control. Geneva. Switzerland: World Health Organization. 2011
- Guo, W. Wang, J. & Wang, S. (2019). Deep multimodal representation learning: a survey. *IEEE Access* **7**, 63373–63394 (2019)
- Giorgio, Q. Ramy, A. Michael, H. & Rima, A. (2021). Machine Learning and the Future of Cardiovascular Care. *Journal of The American College of Cardiology* Vol. 77 , No. 3 ,202 Elsevier on Behalf of the American College of Cardiology Foundation. This IS An Openaccess Article under The CCBY-Ncnd License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).
- Hana, H. & Manal, S. (2021). Detection of Cardiovascular Disease using Machine Learning Classification Models. *International Journal of Engineering Research & Technology (IJERT)* <http://www.ijert.org> ISSN: 2278-0181

- Hung, C.Y., Lin, C.H., Chang, C.S., Li, J.L. & Lee, C.C. (2019). Predicting gastrointestinal bleeding events from multimodal in hospital electronic health records using deep fusion networks.” 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Germany.
- Huang, G. Liu, Z. Van Der Maaten, L., & Weinberger, K.Q. (2017). Densely connected convolutional networks." 2017 IEEE Conf. Computer Vision and Pattern Recognition (CVPR), USA
- Hossam, M. (2019). Cardiovascular Disease Diagnosis: A Machine Learning Interpretation Approach. (*IJACSA International Journal of Advanced Computer Science and Applications*, Vol. 10, No. 12, 2019
- Jane, R. Saravanan, P. Subha P. Arun, R. & Selvaprabu, J. (2020). Prediction of Cardiac Disease using Supervised Machine Learning Algorithms. *Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020) IEEE Xplore* Part Number:CFP20K74-ART; ISBN: 978-1-7281-4876-2
- Keerthana, T. K. (2017). Heart Disease Prediction System using Data Mining Method”, *International Journal of Engineering Trends and Technology (IJETT)*– Volume 47 Number 6 May 2017
- Kunanets, N. Vasiuta, O. & Boiko, N. (2019). Advanced Technologies of Big Data Research in Distributed Information Systems”, in: *Proceedings of the 14th International conference "Computer sciences and Information technologies" (CSIT 2019)*, Lviv, Ukraine, September 17- 20 (2019): 71-76. DOI: 10.1109/STC CSIT.2019.8929756.
- Kamal Kant, and Dr. Kanwal Garg,2014, “Review of Heart Disease Prediction using Data Mining Classifications”, *International Journal for Scientific Research & Development(IJSRD)*, Vol. 2, Issue 04, ISSN (online): 2321- 0613, pp. 109-111
- Kohli, S. (2019, November 18). Understanding a classification report for your machine learning model. Medium. Retrieved November 3, 2022, from <https://medium.com/@kohlishivam5522/understanding-a-classification-report-for-your-machine-learning-model-88815e2ce397>
- Khan Mamun, M.M.R.; Elfouly, T. Detection of Cardiovascular Disease from Clinical Parameters Using a One-Dimensional Convolutional Neural Network. *Bioengineering* **2023**, 10, 796. <https://doi.org/10.3390/bioengineering10070796>
- Latha, C. B., & Jeeva, S. C. (2019). Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques. *Informatics in Medicine Unlocked*, 16, 100203. <https://doi.org/10.1016/j.imu.2019.100203>
- Lee, C. H., & Yoon, H.-J. (2017). Medical Big Data: Promise and challenges. *Kidney Research and Clinical Practice*, 36(1), 3–11. <https://doi.org/10.23876/j.krcp.2017.36.1.3>
- Li, M., Fei, Z., Zeng, M., Wu, F.-X., Li, Y., Pan, Y., Wang, J. (2019). Automated ICD-9 coding via a deep learning approach. *IEEE/ ACM Trans. Comput. Biol. Bioinf.* **16**(4), 1193–1202
- Lu, L., & Harrison, A.P. (2018). Deep medical image computing in preventive and precision medicine. *IEEE Multimedia* **25**(3), 109–113

- Mohammad, A. M. Abushariah, Assal A. M. Alqudah, Omar Y. Adwan, Rana M. M. Yousef, “Automatic Heart Disease Diagnosis System Based on Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) Approaches”, *Journal of Software Engineering and Applications*, 2014, 7, 1055-1064, November 2014 , <http://dx.doi.org/10.4236/jsea.2014>. [26] Soumonos Mukherjee, Anshul Sharma, “Intelligent
- Mohan, S. Thirumalai, C. & Srivastava, G.(2019). Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques. *Access IEEE*. 2019;7:81542–54.
- Nabaouia, L. Meryem, A. & Bouabid, E. (2019). Identification of Cardiovascular Diseases Using Machine Learning. *IEEE Access*
- Nagavelli, U., Samanta, D., & Chakraborty, P. (2022). Machine learning technology-based heart disease detection models. *Journal of Healthcare Engineering*, 2022, 1–9. <https://doi.org/10.1155/2022/7351061>
- Nidhi Bhatla, and Kiran Jyoti, Oct. 2012, An Analysis of Heart Disease Prediction using Different Data Mining Techniques, *International Journal of Engineering Research & Technology (IJERT)*, Vol. 1, Issue 8, ISSN: 2278- 0181, pp. 1-4.
- Patel, J. Tejal, Upadhyay, D. & Patel, S. (2015). Heart disease prediction using machine learning and data mining technique. *Heart Dis*. 2015;7(1):129–37.
- Ramalingam, V. Dandapath, V, A, & Karthik Raja M. (2018). Heart disease prediction using machine learning techniques: a survey. *International Journal of Engineering & Technology*. 2018;7(2.8):684–687. <https://doi.org/10.14419/ijet.v7i2.8.10557>
- Ruch, S. Sandhya, A. Priyanka, P. Prof. Jaishree, W. (2018). An Intelligent Heart Disease Prediction System Using K-Means Clustering and Naïve Bayes Algorithm”,(IJCSIT) *International Journal of Computer Science and Information Technologies*, Vol. 6 (1) , 637-639, IISN : 0975-9646
- 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., ... Fuster, V. (2020). Global burden of cardiovascular diseases and risk factors, 1990–2019. *Journal of the American College of Cardiology*, 76(25), 2982–3021. <https://doi.org/10.1016/j.jacc.2020.11.010>
- Ravi, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., Yang, G.Z. (2017). Deep learning for health informatics. *IEEE Biomed. Health Inf.* **21**(1), 4–21.
- Rubini, P., Subasini, C., Vanitha, A. Kumaresan, V. GowdhamKumar, S.& Nithya, T.M. (2021). A Cardiovascular Disease Prediction using Machine Learning Algorithms.
- Sharma, V. Yadav, S. and Gupta, M. (2020). Heart disease prediction using machine learning techniques,” in *Proceedings-IEEE 2020 2nd International Conference on Advances in Computing, Communication Control*

- and Networking, ICACCCN 2020, vol. 29no. 3, pp. 177–181, Madrid, Spain, 2020 *Annals of R.S.C.B.*, ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 904 – 912
- Sachin, A. Ahmad, A. Ishan, M. Bina, J. & Xi, C. (2020). Machine Learning Strategy for Gut Microbiome-Based Diagnostic Screening of Cardiovascular Disease. *Bioinformatics and Artificial Intelligence Laboratory, Center for Hypertension and Precision Medicine, Program in Physiological Genomics, Department of Physiology and Pharmacology, University of Toledo College of Medicine and Life Sciences, Toledo, OH.*
- Shweta, G. Aditi, N. & Kiran, A. (2017). Heart Disease Prediction using PCA-KNN in Data Mining”, *IDES joint International conferences on IPC and ARTEE*
- Shailaja, K., Seetharamulu, B., & Jabbar, M. A. (2018). Machine learning in healthcare: A Review. 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA). <https://doi.org/10.1109/iceca.2018.8474918>
- Sun, H., Liu, Z., Wang, G., Lian, W., & Ma, J. (2019). Intelligent Analysis of medical big data based on Deep Learning. *IEEE Access*, 7, 142022–142037. <https://doi.org/10.1109/access.2019.2942937>
- Swathy & Saruladha, K. (2021). A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques, *ICT Express* (2021), doi: <https://doi.org/10.1016/j.ict.2021.08.021>
- Tripoliti, E. E., Papadopoulos, T. G., Karanasiou, G. S., Naka, K. K., & Fotiadis, D. I. (2017). Heart failure: Diagnosis, severity estimation and prediction of adverse events through Machine Learning Techniques. *Computational and Structural Biotechnology Journal*, 15, 26–47. <https://doi.org/10.1016/j.csbj.2016.11.001>
- Vincy, Ch. & Bindu, M.S. (2019). Prediction Analysis of Cardiac Disease using Classification (2019). doi: 10.22214/ijraset.2019.6295.
- Weng, S.F. Reys, J. Kai, J. Garibaldi, J.M. & Qureshi, N. (2017). Can machine learning improve cardiovascular risk prediction using routine clinical data? *PLoS ONE*.2017;12(4):e0174944.
- Yazhini, K. & Loganathan, D. (2019). A state of art approaches on deep learning models in healthcare: an application perspective.” 3rd International Conference on Trends in Electronics and Informatics (ICOEI), India. (2019).
- Yu, Y., Li, M., Liu, L., Li, Y. & Wang, J. (2019). Clinical big data and deep learning: applications, challenges, and future outlooks. *Big Data Min. Anal.* 2(4), 288–305 (2019)