International Journal of Engineering and Information Systems (IJEAIS)
ISSN: 2643-640X
Vol. 9 Issue 6 June - 2025, Pages: 135-139

Implementation of Computer-Aided Medical Decision Support System For The Prediction and Classification Of Heart Disease Using Machine Learning

Mr. B.Vinod1 and Dr.G.Geetha2

1Research Scholar, Department of Electronics and Communication Engineering, AVS Engineering College, Salem-Tamil Nadu, India
b.vinod.ece@gmail.com
2Associate Professor, Department of Information Science and Technology, College of Engineering, Guindy, Anna University, Chennai, India,
ggeetha@annauniv.edu

Abstract*—In this paper, an AI-based computer-aided heart disease diagnosis decision support system has been proposed using clinical data, patient information, and electrocardiogram (ECG) data. The proposed system includes three modules: an ECG processor module that allows cardiologists to process and analyze the different waveforms, a machine learning-based heart disease prediction module based on patient information and clinical data, and a deep learning-based 18 heart conditions multiclass classification module using 12- lead ECG data. A user-friendly user interface has also been developed for ease of use of the proposed techniques. Results: The heart disease prediction module was found to be 100% accurate in predicting heart disease based on clinical and patient information, and the multiclass classification module was 93.27% accurate, on average, in classifying heart conditions based on a 12- lead ECG signal. The ECG processor also provides quick diagnosis by analyzing important ECG waveforms and segments. Conclusion: The proposed system may have the potential for facilitating heart disease diagnosis. The proposed method allows physicians to analyze and predict heart disease easily and early, based on the available resource, improving diagnosis accuracy and treatment planning*

Keywords: Artificial Intelligence, AI, clinical data, diagnosis, ECG signal, heart disease

I. Introduction

Cardiovascular diseases (CVDs) are groups of disorders of the heart and the blood vessels including heart coronary heart disease, cerebrovascular disease, peripheral arterial disease, rheumatic heart disease and other conditions. CVDs are the leading cause of death globally, taking an estimated 17.9 million lives each year and more than 75% of these deaths occur in low- and middle-income countries (LMICs).1 Even though evidence on the national burden of cardiovascular diseases (CVDs) is limited in Ethiopia, according to a systematic review conducted in 2014, the prevalence of CVD ranges from 7.2% to 24%.2 The trend of CVD and mortality attributed to CVD is still increasing in Ethiopia.3,4 The risk factors of CVDs could be lifestyle, age and family history (genetic risk).5–7 The most common behavioral risk factors include smoking, obesity, unhealthy diet, lack of physical exercise, and excessive alcohol consumption.1,7,8 Individuals with behavioral risk factors may experience symptoms such as high blood pressure, high blood glucose, high blood lipids, and being overweight or obese.1,9 Identification of risk factors of CVDs early can help prevent premature deaths. For continuous follow-up, accurate diagnosis and risk stratification of patients using a diagnostic laboratory in combination with radiology techniques plays a significant role.10 The common tests to diagnose CVDs include blood work, electrocardiogram (ECG), ambulatory monitoring, echocardiogram, cardiac CT and MRI, stress test, cardiac catheterization etc. However, accurate diagnosis requires analysis and integration of much laboratory data and patient information. Integrated data analysis through the manual procedure can be complex and time consuming, and also the diagnostic effectiveness is dependent on the physicians' knowledge and experience which may sometimes lead to misdiagnosis. Moreover, in countries with limited resources, the diagnosis and treatment of CVDs is usually difficult, due to the unavailability of diagnostic apparatus, low physician-topatient ratios, shortage of high quality medical expertise and infrastructure, resulting in poor prediction and treatment of heart patients.11,12 This burden can be reduced by Introducing clinical decision support systems that encompass a variety of tools to enhance decision making in healthcare.13,14 Artificial intelligence (AI), which is a simulation of human intelligence in machines that are programmed to mimic human thoughts and actions, has the potential to help clinicians make an informed decision in the diagnosis and management of CVDs by analyzing big data. AI based clinical decision support systems can be developed using traditional machine learning (ML) algorithms (a subset of AI that is used to build AI-driven applications) or deep learning algorithms that use large amounts of data and complex algorithms for model training. To overcome the limitations of the manual diagnosis procedures and make use of the potential of AI for disease prediction, literature has proposed different heart disease predictive machine learning techniques based on Support Vector Machines

(SVM), K-Nearest Neighbor (KNN), Naïve Bayes (NB), and Decision Tree (DT), deep learning models and others.9,10–15 For example, Detrano et al19 have used a logistic regression classification algorithm for heart disease detection and claimed a classification accuracy of 77.1%. Similarly, Kahramanli et al21 proposed a heart disease classification system integrating neural networks with an artificial neural network and claimed an accuracy of 82.4%. Likewise, Tomov et al23 came up with a deep neural network model for heart disease prediction, reporting an accuracy of 99% and 0.98 Matthews Correlation Coefficient (MCC). Ali et al25 proposed an expert system using stacked SVM for the prediction of heart disease and reported a 91.11% classification accuracy. To achieve improved clinical diagnosis, a fusion of multimodal data from ECG, clinical laboratory measurement, patient information, etc. are required. However, many of the automatic health disease diagnosis techniques proposed in the literature are either less accurate, and dependent on clinical data, or medical imaging data or ECG signals alone. The purpose of this research is therefore, to develop an integrated robust tool that allows physicians to analyze ECG signals acquired from patients and get a decision support in the prediction and classification of heart diseases using clinical data, patient information and a standard 12 lead ECG record improved because of controlling the surface waves by using lots of PBG microstrip antenna structure [3]. A crystal array

II. METHODS

The current work presents three heart disease diagnosis decision support modules: (1) heart disease prediction module that predicts the presence of heart disease using structured patient information and clinical data, (2) 12-Lead ECG based cardiac condition/abnormality classification which is designed to identify 18 types of cardiac conditions or abnormalities which indicate heart diseases from the 12-lead ECG record, and (3) ECG processor which is designed to process a single lead ECG record and quantifies the important waveform durations, amplitudes and slope. For the first module, structured patient information (age, gender, history of hypertension, etc.), and streaming clinical data (heart rate, blood pressure, etc.), were first processed and analysed. Then feature fusion of the structured data and streaming data was performed to train and validate a machine learning model for heart disease prediction. For the second module, 12-lead ECG data was first pre-processed for artifact removal and the data was used to train and validate a deep learning model (together with age and gender information) for multi-class classification of 18 cardiac conditions. In the third module, an ECG processor that denoises the signal, extracts the QRS complex and ECG waves, analyzes and calculates the ECG wave's amplitudes, duration, and slope as well as the heart rate was developed. Finally, a user-friendly web-based system was developed for ease of use of the proposed sub-systems. Figure 1 demonstrates the general framework of the proposed system

Data Collection For the heart disease prediction system, a total of 1190 observations containing different attribute information including age, sex, chest pain type, blood pressure, cholesterol in mg/dl, blood sugar, maximum heart rate, etc. were acquired from a publicly available database (University of California Irvine (UCI) Machine Learning Repository)26 which was collected

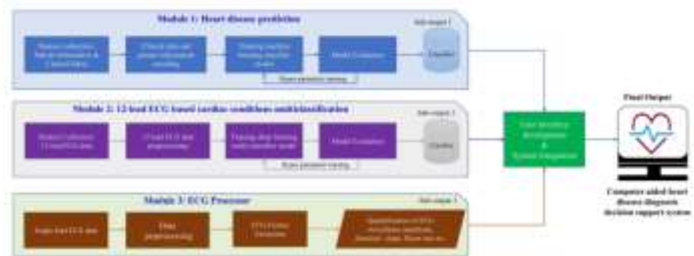

Figure 1 Summary of the proposed computer aided heart disease diagnosis tool

Training and Testing the Heart Disease Prediction Models Two machine learning models (XGBoost and random forest), and an artificial neural network (ANN) deep learning model were trained and tested with the same attribute information for heart disease prediction. The best performing model was then selected for deployment. XGBoost is a decision-tree-based ensemble Machine Learning algorithm that uses a gradient boosting framework. It is an optimized distributed gradient boosting library designed to be highly efficient, flexible and portable It implements machine learning algorithms under the Gradient Boosting framework. XGBoost provides a parallel tree boosting designed to be highly efficient, flexible and portable In this paper, the XGBoost model was implemented with a learning rate of 0.01, L1 regularization value of 5, L2 regularization value of 2, and 2000 number of estimators or runs (model learning iterations). Similarly, random forest is one of the supervised machine learning algorithms that is usually used for both classification and regression purposes. It contains many decision trees that operate as an ensemble. In this paper, the random forest algorithm was implemented with 600 decision trees (estimators) and other default parameters. For training the XGBoost and random forest models, initially, the data was randomly split into training set (80%) and test set (20%). Then, a 10-fold cross validation technique was applied on the training dataset, in which the training seT.

Results Data Pre-Processing and Visualization In the pre-processing stage, the different attribute information used for training the heart disease prediction model were converted into numeric values and analysed. As demonstrated in the correlation plot of Figure 2, chest pain, the maximum heart rate and slope of peak exercise ST segment are highly correlated with the target (having heart disease or not). Figure 4 demonstrates the number of people (in the collected data) with each chest pain type (angina) and the relation between the types of chest pain and heart disease. As indicated, 27.2%

International Journal of Engineering and Information Systems (IJEAIS)
ISSN: 2643-640X
Vol. 9 Issue 6 June - 2025, Pages: 135-139

persons have chest pain type 0, 82% have chest pain type 1, 79.3% have chest pain type 2 and 69.5% have chest pain type 3. As demonstrated in Figure 4, those who have chest pain type 1 and chest pain type 2 are more likely to be affected by heart disease

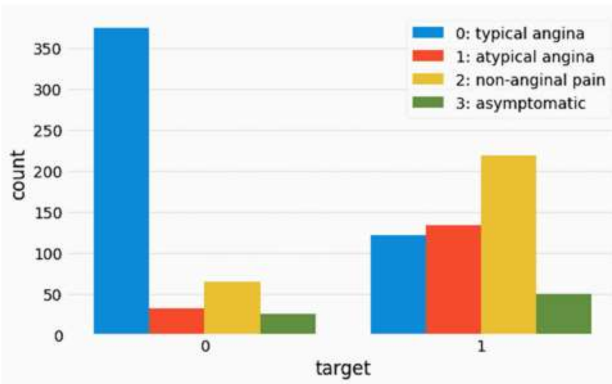

Figure 2 Data visualization demonstrating relation between types of chest pain and heart disease

III. Proposed Design

The training seT was split into 10 parts of equal size, and 9 parts were used for training and 1 part was used for validation. This process was repeated 10 times iteratively and the average of this accuracy was taken as the expected prediction accuracy. The ANN was implemented using a standard feed-forward back-propagation neural network (BPNN) model. The network has three layers, an input layer with 13 neurons, hidden layer with 11 neurons and a 1 neuron output layer. A uniform kernel initializer, ReLu activation function in the input and hidden layer, the sigmoid activation function in the output layer, an Adam optimizer, and a binary cross entropy, batch size of 10 and 100 number of epochs were used in training this model. 80% of the data was used for training and the remaining 20% of the data was used for testing. Training and Testing of the Multiclassification Model For the classification of the 18-cardiac conditions/abnormalities from the 12-lead ECG data, a conventional neural network (CNN) was trained and validated. The model was designed to accept two separate inputs: (i) ECG signal and (ii) age and gender. For the feature extraction of the first input (ECG data), 3 one dimensional conventional neural networks (Conv1D) with 5000 input length and 12 steps were used. For the second input feature extraction two dense layers were used. The outputs of the first and second feature extracting blocks were then concatenated. Finally, a dense layer with 18 outputs was used for final classification. The model uses ReLu activation function for the conventional layers and sigmoid activation function for the dense layer, Adam as an optimizer, and a binary cross entropy loss function. The model was trained for 50 epochs and a batch size of 50. Figure 1 illustrates the simplified architecture of the proposed and the implemented CNN model. ECG Processor An ECG provides key information about the condition of the heart. Analysis based on ECG data is usually conducted after signal processing. ECG data processing techniques include noise removal, baseline correction, wave form and parameter extraction and abnormality detection. An ECG waveform consists of five basic waves called P, Q, R, S, and T-waves and sometimes U-waves. The P-wave indicates the successive depolarization of right atria and left atria, QRS complex indicates the ventricular depolarization, T-wave represents the ventricular repolarization and the U-wave represents the repolarization of the papillary muscles. The most important part of the ECG data analysis is the shape of the QRS complex which is the combination of three of the graphical deflections seen on the typical ECG. Finite impulse response (FIR) digital filters using Kaiser window30 were designed and implemented to remove high frequency noise, low frequency noise, and powerline interference from the ECG data. The low pass and high pass filters were designed with 100 Hz and 0.5 Hz cutoff frequencies, respectively, and order of 100. Similarly, a notch filter with 50 Hz central frequency and order of 100 was designed for removal of the power line interference. After noise removal, ECG feature extraction system was designed to extract important features including R-peak, PQST peaks and waves and each wave amplitudes and intervals. The Neurokit231 discrete wavelet method of ECG peaks detection package was used to extract and delineate the ECG peaks. After extraction of the required peaks, an algorithm was developed for calculation of ST depression, QRS duration, slope of ST segment, QT interval, amplitude of the R peak, amplitude of the Q peak, amplitude of P wave, amplitude of T wave, PR interval, corrected QT interval using Bazett formula32 and the average heart rate. These features are important indicators of the presence of heart disease or abnormality

IV. Simulation Results And Discussions

The heart prediction module accepts attribute information including age, sex, chest pain type, blood pressure, cholesterol level, fasting blood sugar, maximum heart rate, exercise induced angina, ST segment depression, the slope of the peak exercise ST segment, number of major vessels and a blood disorder called thalassemia. Then the system analyses the attributes and predicts whether the person has heart disease or not.

International Journal of Engineering and Information Systems (IJEAIS)
ISSN: 2643-640X
Vol. 9 Issue 6 June - 2025, Pages: 135-139

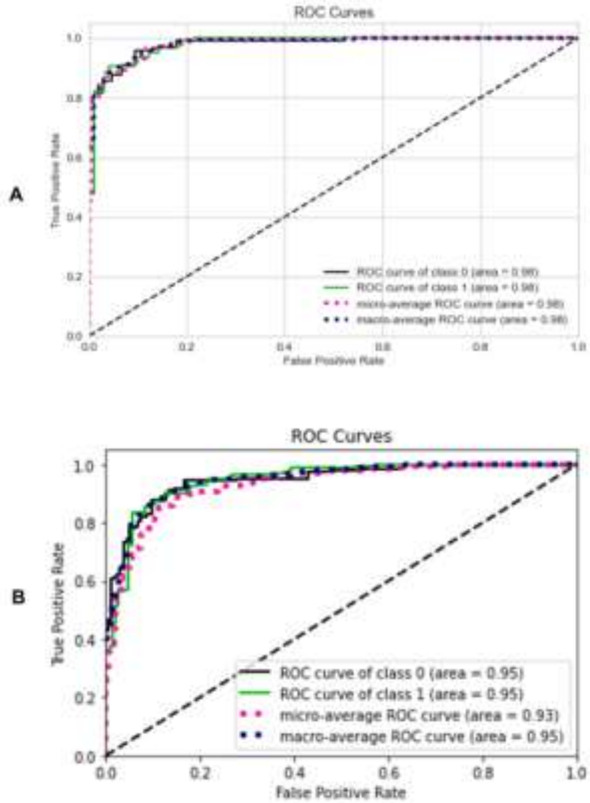

Figure 3 ROC curves of (A) XGBoost and (B) neural network models trained using patient information and clinical data for heart disease prediction

Discussion Heart diseases are the leading cause of death globally. They are fatal diseases that are rapidly increasing in both developed and developing countries. Early diagnosis of heart disease with effective treatment can prevent death or slow down the clinical course. For efficient treatment planning, a variety of tests including laboratory, imaging and other noninvasive techniques are usually required. Physicians usually analyze results of different clinical laboratory tests, visualize and interpret ECG waveforms, measure ECG waveforms' durations/intervals and amplitudes, and integrate all of these results to diagnose the type of heart disease. These traditional manual diagnosis procedures are time consuming, tedious, complex, and dependent on the physician's knowledge and experience which may sometimes lead to misdiagnosis. Due to the limited availability of medical diagnosing tools and medical experts in low-resource settings, diagnosis and cure of heart disease are more complex. Automating the manual diagnosis technique using AI-based predictive techniques could provide quick results helping physicians make informed decisions and reducing diagnosis errors.

V. CONCLUSION

This paper presents an integrated AI-based decision support tool for diagnosis and assessment of cardiac conditions. Different machine learning and deep learning models were trained, evaluated and compared using a variety of data collected from different sources. Best performing models were selected and deployed in a custom designed web-based user interface for the prediction of heart disease and multiclass classification of cardiac conditions. The developed system can provide a reference for clinical diagnosis, remove the opportunities for human error, saves time and money, and improves the diagnosis ability of clinicians for heart disease enabling timely decision making and treatment planning Our experimental results demonstrate that, the developed AI-based computer aided heart disease diagnosis system has the potential to improve diagnostic accuracy, and can be used as a decision support system, especially in those areas where both the means of diagnosis and experts are scarce.

REFERENCES

[1] World Health Organaization. Cardiovascular diseases (CVDs): World Health Organization; 2021 Available from: https://www.who.int/news-room /fact-sheets/detail/cardiovascular-diseases-(cvds). Accessed May 6, 2022.

[2] Misganaw A, Mariam DH, Ali A, Araya T. Epidemiology of major non-communicable diseases in Ethiopia: a systematic review. J Health Popul Nutr. 2014;32(1):1–13

[3] Gebreyes YF, Goshu DY, Geletew TK, et al. Prevalence of high bloodpressure, hyperglycemia, dyslipidemia, metabolic syndrome and their determinants in Ethiopia: evidences from the national NCDs STEPS survey, 2015. PLoS One. 2018;13(5):e0194819. doi:10.1371/journal. pone.019481

[4] Tefera YG, Abegaz TM, Abebe TB, Mekuria AB. The changing trend of cardiovascular disease and its clinical characteristics in Ethiopia: hospital-based observational study. Vasc Health Risk Manag. 2017;13:143–151. doi:10.2147/VHRM.S131259

[5] Hasbani NR, Ligthart S, Brown MR, et al. American Heart Association's life's simple 7: lifestyle recommendations, polygenic risk, and lifetime risk of coronary heart disease. Circulation. 2022;145(11):808–818. doi:10.1161/CIRCULATIONAHA.121.053730.

[6] Knowles JW, Ashley EA. Cardiovascular disease: the rise of the genetic risk score. PLoS Med. 2018;15(3):e1002546–e. doi:10.1371/journal. pmed.1002546.

[7] Said MA, Verweij N, van der Harst P. Associations of combined genetic and lifestyle risks with incident cardiovascular disease and diabetes in the UK Biobank Study. JAMA Cardiol. 2018;3(8):693–702. doi:10.1001/jamacardio.2018.1717 RAECS UIET, Vol. 978, No. 1, pp. 4673-8253, 2015.

[8] Rosiek A, Leksowski K. The risk factors and prevention of cardiovascular disease: the importance of electrocardiogram in the diagnosis and treatment of acute coronary syndrome. Ther Clin Risk Manag. 2016;12:1223–1229. doi:10.2147/TCRM.S107849.

[9] Das R, Turkoglu I, Sengur A. Effective diagnosis of heart disease through neural networks ensembles. Expert Syst Appl. 2009;36(4):7675–7680. doi:10.1016/j.eswa.2008.09.013
[10] Müller MM, Griesmacher A. Rational diagnosis of cardiovascular disease. EJIFCC. 2003;14(2):95–103.
[11] Yang H, Garibaldi JM. A hybrid model for automatic identification of risk factors for heart disease. J Biomed Inform. 2015;58(Suppl):S171–S82. doi:10.1016/j.jbi.2015.09.006.
[12] Coca A, Messerli FH, Benetos A, et al. Predicting stroke risk in hypertensive patients with coronary artery disease: a report from the INVEST. Stroke. 2008;39(2):343–348. doi:10.1161/STROKEAHA.107.495465.
[13] Kiyasseh D, Zhu T, Clifton D. The promise of clinical decision support systems targetting low-resource settings. IEEE Rev Biomed Eng. 2022;15:354–371. doi:10.1109/RBME.2020.3017868
[14] Korom RR, Njue G. Clinical decision support systems in low resource settings. BMJ. 2020;371:m3962. doi:10.1136/bmj.m3962
[15] Patel J, TejalUpadhyay D, Patel S. Heart disease prediction using machine learning and data mining technique. Heart Dis. 2015;7(1):129–137
[16] Haq AU, Li JP, Memon MH, Nazir S, Sun R. A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms. Mob Inform Sys. 2018;2018:3860146. doi:10.1155/2018/3860146
[17] Khourdifi Y, Bahaj M. Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization. Int J Intell Eng Syst. 2019;12:242–252. doi:10.22266/ijies2019.0228.24
[18] Latha CBC, Jeeva SC. Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques. Inform Med Unlocked. 2019;16:100203. doi:10.1016/j.imu.2019.100203
[19] Detrano R, Janosi A, Steinbrunn W, et al. International application of a new probability algorithm for the diagnosis of coronary artery disease. Am J Cardiol. 1989;64(5):304–310. doi:10.1016/0002-9149(89)90524-9
[20] Gudadhe M, Wankhade KK, Dongre SS. Decision support system for heart disease based on support vector machine and artificial neural network. 2010 International Conference on Computer and Communication Technology (ICCCT); 2010:741–745
[21] Kahramanli H, Allahverdi N. Design of a hybrid system for the diabetes and heart diseases. Expert Syst Appl. 2008;35(1–2):82–89. doi:10.1016/j. eswa.2007.06.004
[22] Olaniyi EO, Oyedotun OK, Adnan K. Heart diseases diagnosis using neural networks arbitration. Int J Intell Syst Appl. 2015;7(12):72. doi:10.5815/ ijisa.2015.12.08