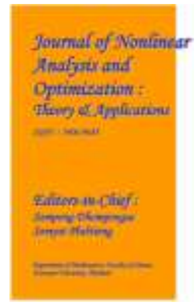
Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 16, Issue. 1: 2025
ISSN : **1906-9685**



DEEP LEARNING APPROACH FOR CARDIOVASCULAR DISEASE DIAGNOSIS USING ANN BASED MULTILAYER PERCEPTRON METHOD

Mr .P.VENKATESWARA REDDY 1, NALLABOTHULA SUMITH VENKAT 2, SHAIK KARISHMA 3, LAKKU HEMA LATHA 4,CHINTHA JOSHITHA 5

#1Assistant Professor in Department of CSE, PBR VISVODAYA INSTITUTE OF TECHNOLOGY AND SCIENCE, KAVALI.

#2#3#4#5 B. Tech with Specialization of Computer Science and Engineering in PBR VISVODAYA INSTITUTE OF TECHNOLOGY AND SCIENCE, KAVALI.

Abstract: The World Health Organisation (WHO) estimates that cardiovascular illnesses claim the lives of 17.7 million people worldwide each year. In order to improve patient outcomes and avoid serious consequences, early identification of cardiac disease is essential. Traditional techniques for predicting heart illness include cardiac MRIs, stress testing, ECGs, and physical exams. These methods, however, are frequently expensive, time-consuming, and reliant on specific medical knowledge.

Artificial Intelligence (AI) and Machine Learning (ML) approaches can be used to uncover useful hidden patterns in the expanding amount of healthcare data. In this study, we use artificial neural networks (ANNs) to develop a system for predicting heart disease. ANNs are strong AI models that can identify intricate links and patterns in medical data, which makes them a great option for predicting cardiac disease.

The Cleveland Heart Disease Dataset from the UCI Machine Learning Repository/Kaggle is used to create the suggested model. Before the ANN is trained using supervised learning techniques, the data is preprocessed. The accuracy and performance of the trained model are then assessed using a different dataset.

The project's ultimate goal is to provide a user-friendly web-based interface that allows patients and medical professionals to enter medical data and get real-time risk assessments for heart disease. Early intervention and improved patient care may be made possible by this system's ability to increase the precision and speed of heart disease detection. The initiative has a major influence on public health

outcomes by advancing AI-driven healthcare solutions.

***Index terms** - Heart Disease Prediction , Artificial Neural Network (ANN) , Machin Learning , Early Detection , Healthcare AI , Cleveland Heart Disease Dataset , Data Preprocessing , Model Training & Testing , Predictive Analytics , Medical Diagnosis*

1. INTRODUCTION

Many people worldwide are affected by coronary sickness, which is a serious general health problem. In many countries, including the US, it is the leading cause of death. Current diagnostic techniques are sometimes costly, time-consuming, and need specialised equipment and knowledge, despite the fact that early detection of cardiac disease is essential for effective treatment and recovery.

By accurately predicting a patient's risk of developing coronary disease based on their clinical history and other relevant factors, artificial intelligence techniques can aid in the early detection of the condition. Of these methods, the capacity of Counterfeit Brain Organisations (ANNs) to predict the likelihood of coronary sickness has demonstrated a high degree of accuracy.

The objective of this job is to develop an ANN model that, given a number of clinical features, can accurately predict a patient's risk of cardiovascular sickness. The model will be trained, tested, and its accuracy and performance assessed using a sizable dataset of patient data.

This project aims to provide healthcare professionals with a reliable and accessible tool for predicting a patient's risk of coronary heart disease. The created model has the potential to help with early identification and treatment of cardiac disease, which can lead to better patient outcomes and better management of resources in the healthcare system. This report's remaining sections are organised as follows: An overview of relevant research in the area of AI-based cardiac disease prediction will be provided in the accompanying part. The project's philosophy and implementation, including dataset preparedness, ANN model development, and model presentation evaluation, will be depicted in the ensuring phases. The conclusion will include a summary of the results and a list of possible improvements to the project.

a) A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques:

https://www.sciencedirect.com/science/article/pii/S2405959521001119

Abstract: Cardio-Vascular Diseases (CVD) are common in the population and cause a high death rate. According to data from a recent poll, the death rate is rising as a result of people's use of tobacco, high blood pressure, cholesterol, and obesity. The aforementioned variables are contributing to the disease's increasing severity. The need of the hour is to research how these variables vary and how they affect CVD. This calls for the use of contemporary methods to detect the illness early on and to help lower the death rate. With their vast array of methods, the fields of artificial intelligence and data mining provide a wide range of research opportunities that might help forecast CVDs in advance and spot their behavioural patterns in vast amounts of data. These forecasts' outcomes will aid physicians in making decisions and identifying patients early, lowering the chance that they will die. The different classification, data mining, machine learning, and deep learning models that are used to predict cardio-vascular disorders are compared and reported in this work. The survey is divided into three sections: Deep Learning Models for CVD Prediction, Machine Learning Models for CVD, and Classification and Data Mining Techniques for CVD. Additionally collated and published in this study are the performance measures used to indicate accuracy, the dataset utilised for classification and prediction, and the tools used for each category of these approaches.

b) Current methods in electrocardiogram characterization:

https://pubmed.ncbi.nlm.nih.gov/24681634/

Abstract: The P-QRS-T wave, which shows the heart's cardiac activity, is called an electrocardiogram (ECG). The sickness affecting the patient is indicated by the little variations in the repolarisation and depolarisation patterns of electric potential. The ECG waveform's clinical time domain characteristics can be utilised to diagnose heart disease. It is quite challenging to correctly identify the ECG classes with the naked eye because of noise and minute morphological parameter values. This study reviews a variety of computer-aided cardiac diagnostic (CACD) systems, analytical techniques, problems solved, and the future of cardiovascular disease screening. The intrinsic differentiating characteristics cannot be adequately represented by techniques like the wavelet transform that were created for time domain, frequency transform domain, and time-frequency domain analysis. Therefore, this paper goes into further depth into nonlinear techniques that may pick up on subtle changes in the ECG signal and offer better accuracy when noise is present. By taking use of these nonlinear characteristics, a CACD system can assist physicians in making more precise diagnoses of cardiovascular illness.

c) Heart disease detection using deep learning methods from imbalanced ECG samples:

https://www.sciencedirect.com/science/article/abs/pii/S1746809421004171

Abstract: Among all illnesses worldwide, heart disease (HD) is the one that kills the greatest number of people. Many important lives will be saved if the illness is detected early and accurately. Computed tomography (CT) images, heart sounds, electrocardiograms (ECGs), and other medical testing can all be used to identify HD. Among all the methods, HD identification from ECG data is essential. The participants' ECG data were regarded in this work as necessary inputs for the HD detection algorithm. Many helpful papers have been published recently that employ various machine learning (ML) and deep learning (DL) models to classify HD. It has been noted that the detection accuracy decreases with unbalanced HD data. In order to improve HD identification, this study has selected appropriate DL and ML models and created and tested the necessary classification models. The goal of the Generative Adversarial Network (GAN) model is to handle unbalanced data by creating and using more fictitious data for detection. Additionally, this work develops an ensemble model that uses GAN and long shortterm memory (LSTM) and performs better than the individual DL model employed in this paper. In comparison to other models, the suggested GAN-LSTM model offers the greatest accuracy, F1-score, and area under the curve (AUC) of 0.992, 0.987, and 0.984, respectively, according to the simulation results using the standard MIT-BIH. Similarly, the GAN-LSTM model performs better than all other models for the PTB-ECG dataset, with accuracy, F1-score, and AUC of 0.994, 0.993, and 0.995, respectively. It is found that the GAN model outperforms the other five models examined, whereas the NB model has the lowest detection potentiality. Additional research may be conducted by selecting all other ensemble models and utilising various datasets, and the results can be produced and compared in a similar manner. Other illnesses and medical issues can also be treated with the suggested optimal diagnostic technique.

d) Cardiac arrhythmia detection using deep learning:

https://www.sciencedirect.com/science/article/pii/S187705091732450X

Abstract: In clinical practice, an electrocardiogram (ECG) is a crucial diagnostic tool for evaluating heart arrhythmias. By categorising patient ECGs into appropriate cardiac states, a deep learning framework that was previously trained on a generic picture data set is applied in this study to perform automated ECG arrhythmia diagnosis. The final classification is performed by feeding the retrieved features into a basic back propagation neural network after a very deep convolutional neural network (AlexNet) has been utilised as a feature extractor. To assess the suggested framework, three distinct ECG waveform circumstances are chosen from the MIT-BIH arrhythmia database. Implementing a straightforward, dependable, and readily adaptable deep learning approach for the categorisation of the three distinct heart diseases that were chosen is the primary goal of this work. The results showed that very high performance rates might be achieved by cascading a traditional back propagation neural network with a transferred deep learning feature extractor. With a testing accuracy of about 92%, the highest recorded accurate identification rate is 98.51%. These findings demonstrated that transferred deep learning was an effective automated strategy for detecting cardiac arrhythmias, removing the need to train a deep convolutional neural network from the ground up and offering a readily implementable solution.

e) HeartID: A Multiresolution Convolutional Neural Network for ECG-Based Biometric Human Identification in Smart Health Applications:

https://ieeexplore.ieee.org/abstract/document/7933065

Abstract: In the new era of smart cities, body area networks—which include smart sensors—are significantly changing the way that health applications are implemented. Physiological signal-based biometric human identification is receiving a lot of attention in order to satisfy growing security and privacy needs. This work focusses on two main obstacles: feature engineering is time-consuming and only fits certain datasets, and signal processing techniques are often both complex and data-dependent. A unique wavelet domain multiresolution convolutional neural network is suggested to provide a highly generalisable and data-independent signal processing and feature learning procedure. In particular, it eliminates the need for the laborious process of extracting signal fiducial properties by enabling the blind selection of a physiological signal segment for identification purposes. The randomly selected signal segment is then converted to the wavelet domain, where multiresolution time-frequency representation is accomplished, in order to enhance the data representation. The phase difference resulting from the blind segmentation procedure is eliminated by applying an auto-correlation technique to the converted data. The human identification job is then carried out by a multiresolution 1-D-convolutional neural network (1-D-CNN), which automatically learns the intrinsic hierarchical features from the wavelet domain raw data without the need for data-dependent and intensive feature engineering. Eight ECG datasets with varying behaviours, including those with or without serious cardiac illnesses, and with various sensor implantation techniques are used to properly assess the efficacy of the suggested approach. Our evaluation has an average identification rate of 93.5% and is far more comprehensive than the state-of-the-art efforts. Even with randomly chosen signal segments and without any feature engineering, the suggested multiresolution 1-D-CNN algorithm is capable of successfully identifying human individuals. It is anticipated that this work will show that using deep learning and blind signal processing techniques for biometric human identification is both feasible and successful. This will allow for a high degree of generalisation with little algorithm engineering effort.

2. METHODOLOGY

i) Proposed Work:

In the suggested approach, we use machine learning, specifically artificial neural networks (ANN), to create the prediction of heart disease. The suggested method is creating an ANN-based prediction model that can precisely identify individuals who are at risk of heart disease. The method employs ANNs to predict the chance of heart illness using the well-known Cleveland Heart illness dataset, which is accessible on Kaggle and the UCI machine learning repository. ANNs are a great option for medical diagnostics since they can handle complicated and noisy data and learn from enormous datasets. constructing an ANN model in the suggested system that can learn from the preprocessed data and precisely forecast a patient's risk of heart disease. An artificial neural network (ANN), often known as a neural network, is a mathematical representation of biological brain networks. The foundation of artificial neural networks is the study of the human brain. The human brain is an intricate network of neurones. Axons, dendrites, and synapses are all found in neurones. The input layer, hidden layer, and output layer are the three layers that make up the planned ANN. The Flask web framework is used in the development of the suggested system, which uses the trained ANN model as a web application to forecast the risk of heart disease in patients and healthcare professionals.

ii) System Architecture:

The proposed heart disease prediction system follows a client-server architecture, where the backend machine learning model is hosted on a Flask-based web server, and the frontend is a user interface that accepts patient details. The system begins with data collection from the Cleveland Heart Disease dataset, followed by data preprocessing such as cleaning, normalization, and transformation. These prepared inputs are fed into an Artificial Neural Network (ANN), specifically a Multi-Layer Perceptron (MLP), which learns patterns and relationships between features like age, cholesterol, blood pressure, and more to predict the risk of heart disease.

Once the model is trained and validated, it is integrated into a Flask application for real-time prediction. When a user enters patient information via the web interface, the data is sent to the server, processed through the trained ANN model, and the prediction result is returned to the user interface instantly. The architecture ensures modularity by separating data handling, model training, and deployment stages. This design allows easy updates, scalability, and integration with medical databases or wearable health monitoring systems in the future.

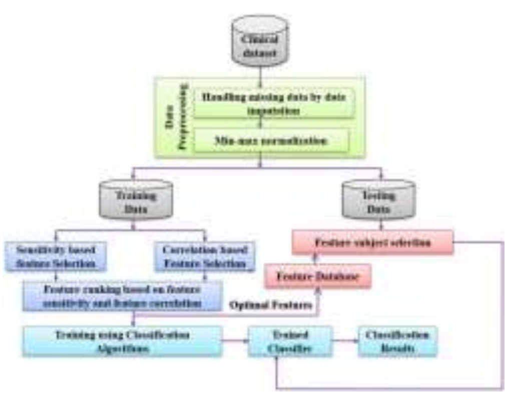


Fig.1 System architecture

iii) MODULES:

a) ***Data Collection Module***

- Collects structured medical data from datasets like Cleveland Heart Disease Dataset.
- Includes key features: age, sex, blood pressure, cholesterol, fasting blood sugar, etc.
- Ensures the dataset has labeled outcomes (0: No Disease, 1: Disease Present).
- May involve integration from multiple health record sources or CSV files.

b) Data Preprocessing Module

- Cleans the dataset by handling missing or inconsistent data values.
- Normalizes data so all features have a similar scale (important for ANN).
- Encodes categorical variables like chest pain type, ECG results using one-hot encoding or label encoding.
- Splits data into training and testing sets (e.g., 80-20 or 70-30).

c) Feature Selection Module

- Uses statistical methods (e.g., correlation matrix) or model-based methods to choose the most relevant features.
- Removes redundant or less significant features to avoid overfitting.
- Improves training speed and overall model performance.

d) Model Training Module

- Builds a Multi-Layer Perceptron (MLP) model using Keras/TensorFlow.
- Consists of input, hidden, and output layers with activation functions (e.g., ReLU, Sigmoid).
- Uses backpropagation and optimizers like Adam for weight adjustment.
- Trained over multiple epochs to minimize error (loss function like binary cross-entropy).

e) Prediction Module

- Accepts new patient input through form or file.
- Processes the input using the same preprocessing steps.
- Predicts the likelihood of cardiovascular disease (binary output).
- Returns confidence score or probability value for medical decision support.

f) Web Interface Module

- Created using Python Flask for real-time user interaction.
- Allows doctors/users to enter patient information via browser.
- Displays prediction output with interpretation (e.g., "High Risk", "Low Risk").
- Can include graphical summary like bar chart or risk meter.

g) Performance Evaluation Module

- Evaluates model accuracy using test data.
- Generates performance metrics: Accuracy, Precision, Recall, F1-Score, ROC-AUC.
- Visualizes results using confusion matrix and ROC curve.
- Helps in tuning hyperparameters for better results.

iv) ALGORITHMS:

Artificial Neural Networks(ANNs)

Comprising connected layers of neurones that learn from data, artificial neural networks (ANNs) are computational models motivated by the human brain. Key components of machine learning, they drive tools including financial forecasting, speech processing, and image recognition.

Often called a neural network, an artificial neural network (ANN) is a computer model motivated by the neuronal organisation of the human brain. Layered organisation defines ANNs, which are made up with linked nodes called "neurones." Every neurone sends information to other neurones, hence allowing the network to understand patterns and make decisions. Especially in jobs requiring pattern recognition, classification, regression, and more sophisticated calculations, artificial neural networks (ANNs) are fundamental to machine learning.

3. EXPERIMENTAL RESULTS

Accuracy: How well a test can differentiate between healthy and sick individuals is a good indicator of its reliability. Compare the number of true positives and negatives to get the reliability of the test. Following mathematical:

Accuracy = TP + TN /(TP + TN + FP + FN)

$$Accuracy = \frac{(TN + TP)}{T}$$

Test Accuracy: 0.9895

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\Pr e\,cision = \frac{TP}{(TP + FP)}$$

Recall: Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a model's completeness in capturing instances of a given class.

$$Recall = \frac{TP}{(FN + TP)}$$

mAP: Mean Average Precision (MAP) is a ranking quality metric. It considers the number of relevant recommendations and their position in the list. MAP at K is calculated as an arithmetic mean of the Average Precision (AP) at K across all users or queries.

$$mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k$$

$$AP_k = \text{ the AP of class } k$$

$$n = \text{ the number of classes}$$

F1-Score: A high F1 score indicates that a machine learning model is accurate. Improving model accuracy by integrating recall and precision. How often a model gets a dataset prediction right is measured by the accuracy statistic.

$$F1 = 2 \cdot \frac{(Recall \cdot \Pr e\,cision)}{(Recall + \Pr e\,cision)}$$

Fig 7 INPUT PARAMETERS

Fig 8 DISEASE PREDICTION PERFORMANCE ANALYSIS

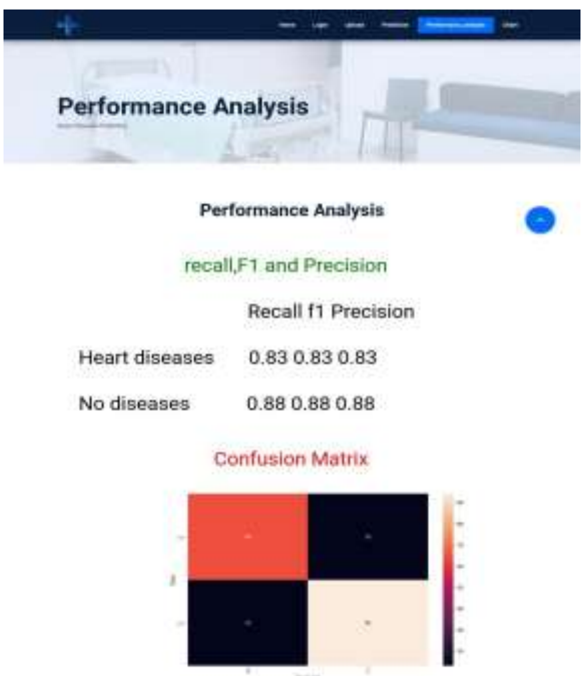


FIG NO:9 PERFORMANCE ANALYSIS DATASET CHART

4. CONCLUSION

To sum up, the study "Heart Disease Prediction using Artificial Neural Network (ANN)" shows how machine learning techniques may be used to effectively forecast a patient's risk of developing heart disease. The project entailed gathering a sizable patient data collection, preparing it, and creating an artificial neural network (ANN) model that can learn from the data and precisely forecast the risk of heart disease. The suggested approach has a number of benefits over current techniques, including high accuracy, scalability, early cardiac disease diagnosis, and customisation. By adding fresh data and updating the model as new information becomes available, the system may be constantly improved. Data collection, dataset preparation, import of required libraries, dataset splitting, ANN model development, model selection, model application, and result analysis were among the components that comprised the project. The created algorithm showed promise in properly forecasting patients' risk of heart disease and attained high accuracy. All things considered, the experiment demonstrates how machine learning methods—more especially, artificial neural networks—can be used to precisely forecast a patient's risk of developing heart disease. Better patient outcomes and more efficient use of healthcare resources might result from this system's ability to help medical practitioners identify and treat cardiac disease early.

5. FUTURE SCOPE

❖ Adding more characteristics: To estimate the risk of heart disease, the existing model employs a collection of features. Nevertheless, additional traits could emerge that might raise the model's accuracy. For instance, combining data from wearable technologies or genetics might yield more thorough information for estimating the risk of heart disease.

- ❖Using more sophisticated ANN models: The present model makes use of a simple ANN model that has several hidden layers. To increase the model's accuracy, more sophisticated ANN models like recurrent neural networks (RNNs) or convolutional neural networks (CNNs) might be employed.
- ❖Adding new data: A comparatively limited dataset is used in the present model. Adding more and more varied datasets might increase the model's precision and resilience.

REFERENCES

[1] Ayatollahi, H., Gholamhosseini, L., & Salehi, M. (2019). "Predicting coronary artery disease: a comparison between two data mining algorithms."

[2] BMC Public Health. doi: 10.1186/S12889-019-6721-5. Benjamin, H., David, F., & Belcy, S. A. (2018). "Heart disease prediction using data mining techniques." ICTACT Journal of Soft Computing, 9(1), 1824-1830

[3] Dengqing Zhang, Yunyi Chen, Yuxuan Chen, Shengyi Ye, Wenyu Cai, Junxue Jiang, Yechuan Xu, Gongfeng Zheng, Ming Chen, "Heart Disease Prediction Based on the Embedded Feature Selection Method and Deep Neural Network", Journal of Healthcare Engineering, vol. 2021, Article ID 6260022, 9 pages, 2021.

[4] Dulhare, U. N. (2018). "Prediction system for heart disease using naïve bayes and particle swarm optimisation." Biomedical Research, 29 (12), 2646-2649.

[5] Gawali, M., & Shirwalkar, N. (2018). "Heart disease prediction system using data mining techniques." International Journal of Pure and Applied mathematics, 120 (6), 499-506.

[6] Haq, A. U., Li, J.-P., Memon, M. H., Nazir, S., & Sun, R. (2018). "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms." Hindawi Mobile Information System. Doi:

10.1155/2018/3860146.

[7] Khan, S. N., Nawi, N. M., Shahzad, A., Ullah, A., & Mushtaq, M. F. (2019). "Comparative analysis for heart disease prediction." International Journal on Informatics Visualization, 1 (4-2), 227-231.

[8] Nashif, S., Raiban, M., Islam, M., & Imam, M. H. (2018). "Heart disease detection by using machine learning algorithms and a real-time cardiovascular health monitoring system." World Journal of Engineering and Technology, 6, 854-873.

[9] Prasead, R., Anjali, P., Adil, S., & Deepa, N. (2019). "Heart disease prediction using logistic regression algorithm using machine learning." International Journel of Engineering and Advanced Technoogy, 8 (3S), 659-662.

[10] Salhi, Dhai Eddine & Tari, Abdelkamel & Kechadi, MTahar. (2021). "Using Machine Learning for Heart Disease Prediction." 10.1007/978-3-030-69418-0_7.

[11] Sharmila, S., & Gandhi, M. P. (2017). "Analysis of heart disease prediction using data mining techniques." International Journal of Advanced Networking and Applications, 8 (5), 93- 95.

[12] Sridhar, A., & Kapardhi, A. (2018). "Predicting heart disease using machine learning algorithms." International Research Journal of Engineering and technology, 6 (4), 36-38