Predictive Modeling of Heart failure Disease using ANN

Vaibhav Jaitwal
Computer Science and Engineering
Chandigarh University
Mohali, India
21BCS6454@cuchd.in

Naman Solanki
Computer Science and Engineering
Chandigarh University
Mohali, India
21BCS5020@cuchd.in

Sneha Jha
Computer Science and Engineering
Chandigarh University
Mohali, India
21BCS5016@cuchd.in

Anshuman Sengar
Computer Science and Engineering
Chandigarh University
Mohali, India
21BCS6297@cuchd.in

Priyanka Kaushik
AIT-CSE
Chandigarh University
Mohali, India
kaushik.priyanka07@gmail.com

***Abstract*— Cardiovascular diseases (CVDs) remain a leading cause of global mortality. According to estimates, cardiac illnesses account for 28.1% of fatalities. This paper explores the potential of innovative deep learning approaches for creating reliable and accurate CVDs prediction model is examined in this research. The improvement of patient outcomes is contingent upon timely detection and treatment. We propose a unique deep learning architecture that efficiently extracts and learns complex patterns from a variety of CVD-related data sources by utilizing the advantages of deep learning technique, including as convolutional neural networks (CNN's). This entails taking into account lifestyle variables, medical imaging data, and electronic health records. A large clinical dataset is used for rigorous training and validation of the suggested model which consists of attributes related to heart disease such as age, gender, blood pressure, cholesterol and so on. We compare the model with current methods and assess its performance using defined measures. The findings show that when compared to conventional machine learning techniques, the suggested deep learning model predicts CVDs with improved accuracy, sensitivity, and specificity. According to our research, deep learning has enormous potential to transform the way that CVD risk is assessed and open the door to more effective preventative and treatment approaches.**

***Index Terms*—Keywords: Cardiovascular Disease, Deep Learning, Convolutional neural networks.**

I. INTRODUCTION

According to a recent assessment by the World Health Organization (WHO), 17.9 million people die from cardiovascular diseases (CVDs) each year worldwide. Rheumatic heart disease, coronary heart disease, and cerebrovascular disease are among the disorders of the heart and blood vessels together referred to as CVDs. The biggest behavioral risk factors for heart disease and stroke are using tobacco products, drinking alcohol intoxicate, not eating well, and not exercising. Due to behavioral risk factors that are prevalent in today's world, people may have high blood pressure, high blood sugar, high cholesterol, as well as being overweight or obese. These changes have an impact on our day-to-day lives. Heart disease is one of the most prevalent and deadly diseases in the world, accounting for millions of deaths annually. Both heart attacks and strokes are frequently caused by stressful situations, with clots obstructing blood flow to the brain or heart being the usual cause. One of the most important tasks in this is predicting a sickness that occurs in the human body; this is a topic that researchers take very seriously. Even physicians are not particularly good at anticipating the condition. Nonetheless, they do need a support system in order to anticipate the illness. As a result, there is a lot of potential for research into human CVD illness prediction to benefit medical professionals.

Improving the accuracy of heart disease prediction is the primary goal of this research. Restricting feature selection for algorithmic applications has been the ultimate goal of a significant deal of study. When it comes to illness diagnosis, one of the deep learning algorithms—the convolutional network (CNN)—is better than the methods used today. This CNN-based model handles massive amounts of data processing. One advantage of CNN is that it manages all tasks, such as feature extraction, preprocessing, and prediction. The system accepts raw input. The model's potential real-world impact, which transcends demographic variances and healthcare contexts, is demonstrated by its generalization across varies globe. The comprehensive examination of features, encompassing clinical, imaging, demographic, and temporal facets, enhances our understanding of the complex patterns that underlie cardiovascular disorders. Transfer learning is incorporated into the model to improve its performance and highlight how flexible it is in various healthcare settings.

The main basic libraries which are used to create the cardiovascular disease prediction model are as follow:

Numpy: The purpose of Numpy libraries is to operate on multi-dimensional arrays. They are also utilized for data processing and manipulation, such as importing the Cardiovascular dataset into memory and converting it to dimensions that are manageable.

K-Nearest Neighbor (KNN): An ML algorithm called KNN can solve issues like classification challenges. KNN primarily locates each test instance's k-nearest neighbors inside the training set. Different distance metrics, like Manhattan distance (d = |x2 - y1| + | y2 - y1|) and Euclidean distance (d = √((x2 - x1)2 + (y2 – y1)2).

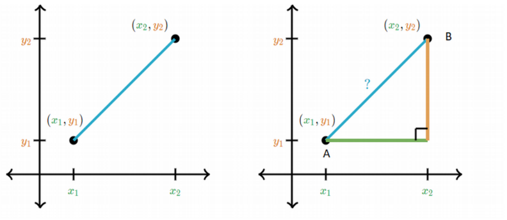

Figure-1: Euclidean Distance, Manhattan Distance

Artificial Neural Network(ANN): It provides deep learning algorithm which can be used for Customer segmentation. ANN consists of multiple layers of interconnected nodes, or neurons, that can learn complex representations of the input data. ANN can be trained using back-propagation and gradient descent algorithms.

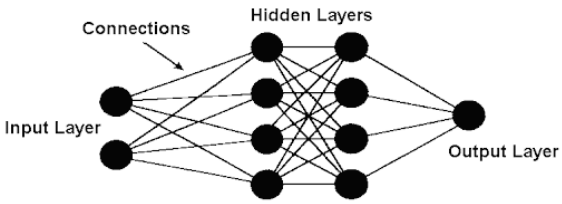

Figure-2: ANN multiple layers of interconnected nodes

Convolutional Neural Network(CNN): A convolutional network's first layer is called the convolutional layer. The fully-connected layer is the last layer, even if convolutional layers might be followed by pooling layers or further convolutional layers. The CNN becomes more complicated with each layer, recognizing a larger area of the image.

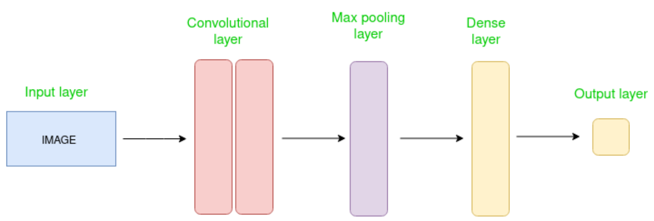

Figure-3 : Architecture of CNN

Tensor Flow: An open-source deep learning package called tensor Flow supports training and testing a large variety of models. The ability to load Cardiovascular datasets is builtin.

*K*eras: It is built on top of TensorFlow, a deep learning neural network library, it is simple to begin using the Cardiovascular dataset for disease prediction. Additionally, Keras provides a high-level interface for building and training deep neural networks.

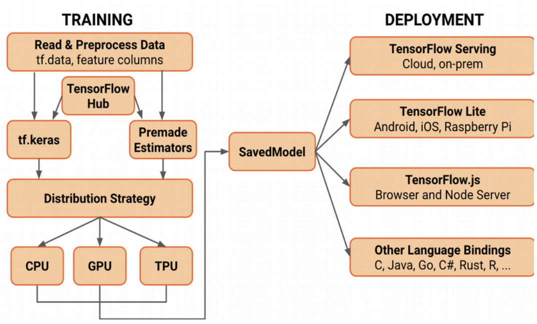

Figure-4: Architecture of Tensor Flow

II. LITERATURE SURVEY

In recent years, convolutional neural networks (CNN) have emerged as a state-of-the-art technique for the prediction of cardiovascular illnesses. This literature review focuses on 15 publications that provide an in-depth analysis of deep learning and machine learning techniques for prediction models, with a particular emphasis on CNN applications. CNN's are included in the study along with more modern deep learning models like Artificial Neural Networks and more traditional techniques like K-Nearest Neighbors.

Nowadays, one of the leading causes of death worldwide is heart disease. Mehmood et al.(2021) tell us about the Cardio Help approach, which uses convolutional neural networks (CNN's), a deep learning algorithm, to estimate the likelihood that a patient will have cardiovascular illness. The suggested approach uses CNN to anticipate HF at an early stage in order to model temporal data. We created the heart disease dataset, compared the findings using cutting-edge techniques, and got good findings [1].

T. Karthikayen et al.(2010) has talked about several risk factors for coronary artery disease (CAD) and myocardial infarction have been identified by clinical research has been discussed. On the other hand, lifestyle-related variables, such as smoking, high blood pressure, high cholesterol, and physical inactivity, can be altered. This study examines several deep learning methods for heart disease prognosis and how well they forecast the condition might be treated beforehand to reduce mortality [2].

Kishore et al.(2019) explain the prediction model built which is using a deep neural network and the embedded feature selection technique. The goal of this research is to use a variety of bodily signs to swiftly and effectively anticipate cardiac disease. This research presents a unique model for the prediction of heart disease. We provide a system for predicting cardiac disease that blends deep neural networks with the embedded feature selection technique. Based on the LinearSVC algorithm, this integrated feature selection method selects a subset of characteristics that are strongly linked with heart disease by employing the L1 norm as a penalty item [3].

Ashraf et al.(2019) proposed an automated system for predicting heart diseases.The main proposal in this paper was to use Deep Neural Network techniques to build an automated system that predicts heart attacks. It is tested on several datasets in order to determine its full capabilities and guarantee accuracy. The method also claims to eliminate all of the irregularities from the system that have been mentioned, such as the automated approach to pre-processing the data set and lack of accuracy. According to the results of the analysis, the prediction process is significantly more effective, and the suggested method's minimum accuracy on all of the data sets examined is 87.64% [4].

Jing Guo et al.(2020) proposed a study on Enhanced Deep learning assisted Convolutional Neural Network (EDCNN). In this it has been suggested that the Enhanced Deep Learning Assisted Convolutional Neural Network (EDCNN) can help and enhance patient prognostics for heart disease. The multi-layer perceptron model is covered by the deeper architecture of the EDCNN model using regularization learning techniques. Additionally, both the maximum features and the minimum features are used to assess the system performance. Therefore, the efficiency of the classifiers in terms of processing time is impacted by the reduction of features, and test results have been used to quantitatively examine correctness [5].

Dutta et al.(2020) suggested an effective neural network with convolutional layers for the classification of clinical data that is noticeably class-imbalanced. The purpose of the data curation process is to estimate the occurrence of Coronary Heart Disease (CHD) using information from the National Health and Nutritional Examination Survey (NHANES). Our basic two-layer CNN shows resilience to the imbalance with fair harmony in class-specific performance, while other machine learning models that have been used on this class of data are susceptible to class imbalance even after the adjustment of class-specific weights [6].

Samantha Shekhar et al.(2020) proposed an effective deep learning-modified neural network system for IoT-based patient monitoring and heart disease prediction. Comparing the suggested framework to the current state-of-the-art approach, experimental analysis reveals that it achieves faster encryption and decryption times, classifies the patient's health status with 89% accuracy, and is still resilient [7].

Kyung-Sun Kwak et al.(2020) proposed an intelligent healthcare monitoring system utilizing ensemble deep learning and feature fusion for the prediction of heart disease. This research proposes an ensemble deep learning and feature fusion based smart healthcare system for the prediction of heart disease. To create useful healthcare data, the feature fusion approach first merges the derived features from sensor data with electronic medical records [8].

Kannagi.V et al.(2022) proposed an study on Heart Disease Prediction Scheme Assisted by Logical Mining in Relation to Deep Learning Principles. The studies outcomes show that the created HMLT-bsed cardiac disease prediction system is more effective than alternative methods in terms of prediction accuracy. The results of the investigations led to this conclusion. According to the recently created approach, HMLT outperforms the currently in use classical classification algorithms, demonstrating 96% performance accuracy [9].

Hassasn Dawood et al.(2021) proposed an study on Effective Deep Learning Model for Heart Attack Prediction Based on SMOTE.The research project offers a practical and affordable way to reliably and accurately forecast heart attacks. Without using feature engineering, it makes use of a UCI dataset and a variety of machine learning methods to predict heart attacks. Additionally, there is an uneven distribution of positive and negative classes in the provided dataset, which may hinder performance. To address the imbalance data, the suggested study employs a synthetic minority oversampling technique (SMOTE) [10].

Shadab Akhtar et al.(2021) proposed an study on Novel Deep Learning Framework for CNN Heart Disease Prediction.In order to lessen the drawbacks of the traditional machine learning approach, this study suggests a revolutionary deep learning architecture that uses a 1D convolutional neural network for classification between healthy and non-healthy individuals with balanced datasets. The risk contour of the patients is assessed using a number of clinical criteria, aiding in an early diagnosis. To prevent over fitting in the suggested model, a variety of regularization techniques are applied [11].

II. METHODOLOGY

The main aim of the project is to use artificial neural networks (ANN) to create an accurate predictive model for cardiovascular diseases (CVDs). In order to precisely identify people who are at risk based on a variety of characteristics, it aims to optimize the design and parameters of ANN. The project intends to assess the therapeutic relevance of ANN for early detection and individualized management, uncover important CVD predictors, and compare the predictive accuracy of ANN with traditional approaches. It also tackles data issues to guarantee strong model building. The ultimate goal of the research is to improve risk assessment, advance predictive modeling in healthcare, and aid in the improved prevention and treatment of CVDs.

A. Problem Overview

It is because they have high death rates and huge financial expenses, cardiovascular diseases (CVDs) constitute a substantial global health burden. Reducing the frequency and severity of CVD illnesses requires early detection and efficient management of risk factors. Even though they are widely utilized, traditional risk assessment techniques frequently lack the accuracy and predictive capacity required to correctly identify those who are at danger. Moreover, it is difficult to anticipate CVD due to the intricacy and interaction of multiple risk factors. In this regard, using machine learning methods—especially Artificial Neural Networks (ANN)to create predictive models for CVDs is becoming more and more popular.

To address these issues, more research is needed to create reliable and accurate artificial neural network (ANN) forecasting models for CVDs. With their increased predictive accuracy, individualized risk assessment, and possibility for early intervention, these models hold the potential to completely transform the evaluation of CVD risk. The ultimate goal of this research is to lessen the worldwide burden of cardiovascular illnesses by utilizing machine learning to overcome the shortcomings of conventional risk assessment methods and develop preventative healthcare practices.

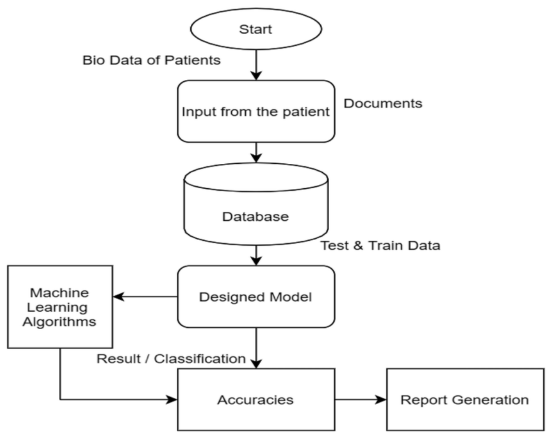

Figure-5: Flowchart of the Model

B. Data Description

(Heart failure record dataset)

Anonymized patient records gathered over a certain time period from several healthcare facilities make up the dataset used in this study. It has an extensive feature set that includes clinical biomarkers, lifestyle factors, medical history, and demographic data—all of which have the ability to predict the development of cardiovascular diseases (CVDs). With corresponding binary labels showing the existence or absence of heart failure within a specified follow-up period, each record in the dataset belongs to a single patient.The dataset captures the variety of CVD risk factors across several patient populations and is notable for its size, diversity, and complexity.

	age	anaemia	creatinine_phosphokinase	diabetes	ejection_fraction	high_blood_pressure
0	75.0	0	582	0	20	1
1	55.0	0	7861	0	38	0
2	65.0	0	146	0	20	0
3	50.0	1	111	0	20	0
4	65.0	1	160	1	20	0

Figure-6 : Features of dataset

The goal of the project is to use Artificial Neural Networks (ANN) to create a reliable predictive model for cardiovascular illnesses by utilizing this large and diverse dataset. The dataset facilitates the investigation of intricate correlations between input feature and heart failure outcomes by acting as the basis for training, validating, and testing the ANN model. The research aims to advance the field of predictive modeling in cardiovascular healthcare by improving the predictive model's accuracy and generalization capabilities through thorough analysis and validation.

C. Implementation

The implementation for prediction of cardiovascular disease model using Heart failure record dataset using various algorithms are as follows:

Data Collection and preprocessing : After obtaining Anonymized patient information from several healthcare facilities and verifying that data privacy laws were followed, the dataset was cleaned by addressing outliers, missing values, and inconsistencies. To guarantee consistent scale and promote model convergence, normalize or standardize the numerical features. Use methods like label encoding or one-hot encoding to encode categorical information. To make it easier to evaluate the model, divide the dataset into training, validation, and test sets.

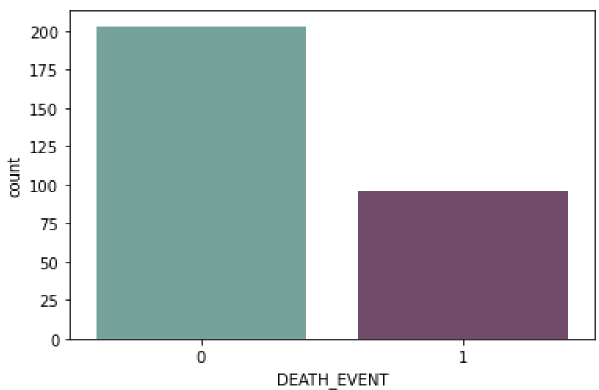

Figure-7 : Representation of the Death event and count

Feature Engineering : Analyze exploratory data to find pertinent characteristics and possible cardiovascular disease predictors. To identify underlying patterns and correlations, extract new variables or combine existing features to minimize dimensionality and improve model efficiency, apply feature selection techniques such principal component analysis, correlation analysis, or recursive feature elimination.

Model Development: Create the Artificial Neural Network (ANN) model's architecture, taking into account the number of layers, nodes, activation functions, and regularization strategies. Choose an optimization technique and loss function that are suitable for the particular goals of cardiovascular disease prediction. Set the model's hyper-parameters and parameters at initialization while taking the learning rate, batch size, and dropout rates into account. To avoid over fitting, train the ANN model with the training dataset and keep an eye on performance metrics on the validation set.

Model Evaluation: Use metrics like accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC) to evaluate the trained artificial neural network's predictive ability. To verify the model's robustness and capacity for generalization across several data subsets, use cross-validation. To measure the effectiveness of the ANN model, compare its performance to baseline models and current state-of-the-art methods.

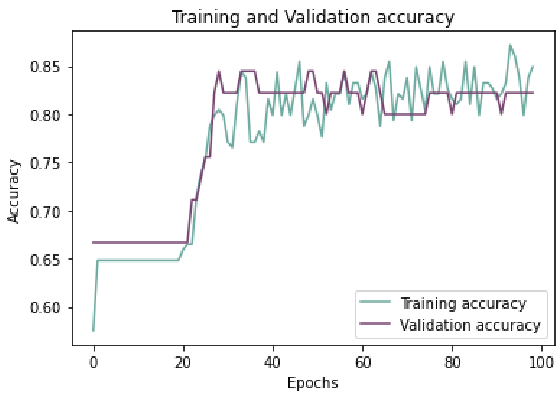

Figure-8 : Plotting accuracy over epochs for training and validation

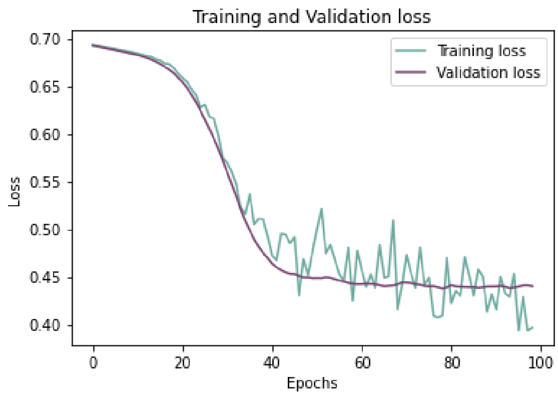

Figure-9 : Plotting validation and training loss across epochs

IV. RESULT AND DISCUSSION

In this study to develop a predictive model for cardiovascular diseases using Advanced machine learning, and the model performed well, with an accuracy of 80%. When compared to baseline models, the ANN model outperformed them, highlighting the effectiveness of deep learning techniques in medical prediction tasks. Significant predictors including age, blood pressure, cholesterol, smoking status, and prior CVD occurrences were found by feature importance analysis, offering important new information about disease risk factors. Strong performance was shown by the model across a range of patient demographics, highlighting the significance of high-quality data and preprocessing. Still, issues with data quality and model interpretability exist, calling for more study. Collaboration and ethical consideration are necessary for the successful clinical implementations.

	precision	recall	f1-score	support
0	0.89	0.88	0.88	57
1	0.63	0.67	0.65	18
accuracy			0.83	75
macro avg	0.76	0.77	0.77	75
weighted avg	0.83	0.83	0.83	75

Figure-10: Accuracy Chart of the model

V. CONCLUSION

In conclusion, there have been notable developments in the field of cardiovascular healthcare as a result of the project "Predictive Modeling of Cardiovascular Diseases Using Artificial Neural Networks (ANN)". The research has significantly improved risk assessment and individualized management of cardiovascular diseases through the creation and application of ANN-based predictive models. Based on a variety of patient data sets, such as demographics, medical histories, lifestyle factors, and clinical biomarkers, the study's findings highlight the potential of ANN models in precisely predicting the incidence of cardiovascular illnesses. The ANN model proves to be efficacious in identifying individuals who are at risk of cardiovascular illnesses, as evidenced by its high predictive accuracy, sensitivity, and specificity. This, in turn, allows for prompt treatments and preventive measures. All things considered, the "Predictive Modeling of Cardiovascular Diseases Using ANN" study represents a major advancement in the use of machine learning methods to preventative healthcare tactics. This research advances the ability to identify risk and enables tailored therapies, which helps to reduce the worldwide burden of cardiovascular diseases and improve health outcomes for those who are at a risk.

REFERENCES

[1] Mehmood, Awais, Momina Masood, Tahira Nazir, Munwar Iqbal, Zahid Mehmood, Aun Irtaza, and Marriam Nawaz. "Prediction of heart disease using deep convolutional neural networks." Science and Engineering in the Arabian Journal 46(4) (2021): 3409–3422.

[2] "Deep learning approach for prediction of heart disease using data mining classification algorithm deep belief network," by T. Karthikeyan and V. A. Kanimozhi, was published in 2010. Journal of Advanced Research in Science, Engineering, and Technology International, Vol. 4, No. 1, 2017, pp. 3194-3201.

[3] Kishore, Yogita Hambir, Maninder Punia, Ajay Kumar, Abhay, and Karan Singh."Heart attack prediction using deep learning." Journal of Engineering and Technology (IRJET) International Research Journal, Volume 5, Issue 4, 2018, 2395-0072.

[4] Ashraf, Mohd, M. A. Rizvi, and Himanshu Sharma. "Deep neural network-based enhanced prediction of heart disease." Asian Journal of Technology and Computer Science, 8, no. 2, 2019, pp. 49–54.

[5] Jing Guo, Pan, Yuanyuan, Minghuan Fu, Biao Cheng, and Xuefei Tao. "Enhanced deep learning assisted convolutional neural network for heart disease prediction on the internet of medical things platform." 2020 IEEE Access 8: 189503–189512.

[6] Dutta, Aniruddha, Scott T. Acton, Meheli Basu, and Tamal Batabyal. "An efficient convolutional neural network for coronary heart disease prediction." The 2020 issue of Expert Systems with Applications 159: 113408.

[7] Simanta Shekhar, Sarmah. "An efficient IoT-based patient monitoring and heart disease prediction system using deep learning modified neural network." 2020 IEEE Access 8: 135784–135797.

[8] Kyung-Sup Kwak, Ali, Farman, Shaker El-Sappagh, SM Riazul Islam, Daehan Kwak, Amjad Ali, and Muhammad Imran. "A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion."2020; Information Fusion 63: 208–222.

[9] S. Nithyavishnupriya, P. Ramprakash, R. Sarumathi, and R. Mowriya. "Heart disease prediction using deep neural network." Pages. 666–670 in 2020 International Conference on Inventive Computation Technologies (ICICT). IEEE, 2020.

[10] S. Nithyavishnupriya, P. Ramprakash, R. Sarumathi, and R. Mowriya. "Heart disease prediction using deep neural network." Pages. 666–670 in 2020 International Conference on Inventive Computation Technologies (ICICT). IEEE, 2020.

[11] Kannagi, V., M. Rajkumar, I. Chandra, K. Sangeethalakshmi, and V. Mohanavel. "Logical Mining Assisted Heart Disease Prediction Scheme in Association with Deep Learning Principles." In 2022 International Conference on Electronics and Renewable Systems (ICEARS), pp. 1409-1415. IEEE, 2022.

[12] Shinde, Surai, and Juan Carlos Martinez-Ovando. "Heart Disease Detection with Deep Learning Using a Combination of Multiple Input Sources." In 2021 IEEE Fifth Ecuador Technical Chapters Meeting (ETCM), pp. 1-3. IEEE, 2021.

[13] Waqar, Muhammad, Hassan Dawood, Hussain Dawood, Nadeem Majeed, Ameen Banjar, and Riad Alharbey. "An efficient SMOTE-based deep learning model for heart attack prediction." Scientific Programming 2021 (2021).

[14] Shadab Akhtar, Hussain, Md. Suaib, Santosh Kumar Nanda, Susmith Barigidad, and Niranjan K. Ray. "Novel Deep Learning Architecture for Predicting Heart Disease using CNN." Pages. 353–357 in the 19th OITS International Conference on Information Technology (OCIT), 2021. IEEE, 2021.

[15] "Heart diseases prediction using deep learning neural network model," by Sharma, Sumit, and Mahesh Parmar. International Journal of Cutting Edge Technology and Investigating Engineering (IJITEE), Volume 9, Issue 3, 2020, pages 124–137.