IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

3rd-International Conference on Muti-Disciplinary Application & Research Technologies (ICMART-2024)

Geetanjali Institute of Technical Studies

Vol. 11, Special Issue 2, May 2024

# A Review: Cardiovascular Disease Prediction Model using Machine Learning and Deep Learning

**Charvi Bapna[1], Charu Kavadia[2], Umang Pahuja[3], Mehul Pokharna[4]**

UG Scholars, Department of Computer Science and Engineering, Geetanjali Institute of Technical Studies, Dabok, Udaipur, Rajasthan, India [1,3,4]

Assistant Professor, Department of Computer Science and Engineering, Geetanjali Institute of Technical Studies, Dabok, Udaipur, Rajasthan, India [2]

**Abstract***:* The whole world has come to know the fact that heart disease is not a trivial issue. Although the years have changed, throngs of patients are diagnosed with this lethal disease and not only is it not decreasing, but increasing and this is evident from the analysis of death rates across the country. Heart disease is caused by several main factors such as negligence in taking care of diet and daily life. Causes such as age and genetics cannot be fully controlled by humans. So it is the responsibility of each individual to take care of critical risk factors that can be controlled by humans. Innovation and data mining technology should also help to reduce the death rate caused by heart disease. Existing data should be used to find hidden relationships, which are very useful for early diagnosis, which can be a new advanced technology and potentially improve the level of the national health system. This study highlights earlier research that uses deep learning and machine learning knowledge to forecast cardiac disease. The accuracy produced by each algorithm varies, and this study will look at the factors that make various outcomes possible. Upon completion of the study, the outcomes attained using deep learning demonstrated superior performance compared to machine learning, exhibiting consistency within a tolerable range of 84% to 99%.

**Keywords***:* Heart disease, Data mining technology, Deep learning, Machine learning, National health system, Death rates

## I. INTRODUCTION

The word "heart disease" is not new, nor is it unfamiliar to people worldwide. It is acknowledged as the primary cause of death worldwide year after year, making it the greatest worry for people of all ages. All disorders of the heart and its activities are included in the category of heart disease. The most frequent condition in the US population to be heard of and experience is coronary artery disease. This is a condition in which an irregularity in the blood supply to the heart might result in a heart attack. Even though heart disease can be deadly, it can be treated if a person can identify it early. Therefore, it should come as no surprise that a medical professional highly advises someone exhibiting heart disease symptoms to get a diagnosis as soon as possible. But there are still holes that make heart disease tough to manage despite the best efforts of the doctors. The application of data mining in healthcare is still not fully understood by a vast number of people. Every medical facility has a vast amount of patient data that might serve as the foundation for risk factor analysis of serious illnesses like heart disease.

Furthermore, some people decide not to give their daily food and lifestyle any thought. Many worldwide health organizations identify high blood pressure, high cholesterol, and smoking habits as the three (3) key risk factors for heart disease. As was previously noted, early detection is highly recommended by medical professionals due to its tremendous potential to allow adequate time for therapy to stop heart disease from getting worse. Regrettably, a certain amount of suspicion regarding early diagnosis persists among medical professionals because it is unclear how effective early diagnosis is, making it seem insignificant.

This study will gather variations of heart disease prediction studies using machine learning and deep learning that have been done from the year until now additionally, we will also see the accuracy obtained till now.

## II. LITERATURE REVIEW

The focus of this literature review is heart disease prediction algorithms that use machine learning and deep learning. Among the studies found are as follows:

Rahman et al.[1] introduced a novel self-attention-based transformer model to forecast the risk of CVD by combining transformer networks with self-attention mechanisms. The Cleveland dataset, a benchmark dataset from the University of California, Irvine (UCI) machine learning (ML) repository, was used to evaluate the suggested model. This model was a helpful tool in the healthcare industry since it could rapidly and easily adapt to various outcome risk prediction and evaluation difficulties.

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

3rd-International Conference on Muti-Disciplinary Application & Research Technologies (ICMART-2024)

Geetanjali Institute of Technical Studies

Vol. 11, Special Issue 2, May 2024

Several medical and non-medical elements in this dataset could be utilized to determine whether a patient had heart disease. The results of the investigation indicated that cardiac disease could be identified with 88.7% accuracy.

Mansoor et al.[2] created a model that could predict CVD illnesses with high accuracy, hence lowering the disease's fatality rate. For this work, they employed a benchmark dataset of 14 distinct heart disease-related parameters from UCI Heart disease prediction. This work explored the application of deep learning models for cardiovascular disease (CVD) prediction using the Cleveland CVD dataset from the UCI Machine Learning Repository. Using the UCI dataset, this work presented an accuracy score summary of CNN, ANN, and LSTM algorithms for heart condition prediction.

Dhaka et al.[3] sought to solve this important problem by presenting a novel hybrid classifier-based disease prediction model. Deep Bidirectional Long-Short-Term Memory (deep Bi LSTM) and deep Convolutional Neural Network (deep CNN) were coupled to create the suggested hybrid classifier. The heart disease dataset used 14 databases and 76 features, while the comprehensive heart disease dataset incorporated separately available datasets based on particular features. This model performed better than earlier models in several ways and employed a modified Diffi-Huffman algorithm for secure data transmission.

Demir et al.[4] wrote in his research about the potential of deep learning for heart disease prediction. It highlighted a gap in existing research, which often relied on machine learning methods. By employing Convolutional Neural Networks (CNNs) and Deep Neural Networks (DNNs) on a heart disease dataset from Kaggle, the study achieved high accuracy (over 93%). This surpassed the performance of some previous research, underlining the potential of deep learning for early detection of this critical health condition.

Sharma et al.[5] proposed a deep learning model for predicting coronary heart disease (CHD). Focused on real-world applications in hospitals, the model aimed to surpass traditional machine learning methods by achieving higher accuracy and reducing false negatives. The research utilized a large and relevant dataset – the BRFSS-2015 (Behavioral Risk Factor Surveillance System) – and demonstrated the proposed model's effectiveness with an impressive accuracy of 98.28%.

Gopalakrishnan et al.[6]emphasized the critical role of early heart disease detection and proposed a Convolutional Neural Network (CNN) model as a valuable tool. CNNs excelled at extracting crucial information from raw medical data, making them suitable for disease prediction tasks. The study demonstrated the effectiveness of the proposed CNN model by achieving improved accuracy, highlighting its potential for earlier and more precise heart disease detection.

Garcí¬a-Ordás et al.[7] tackled the challenge of achieving high accuracy in heart disease prediction with limited data. It proposed a novel approach that combined a Sparse Autoencoder (SAE) and a Convolutional Neural Network (CNN). The dataset used contained only 918 samples with 11 clinical features per patient. By jointly training the SAE and CNN, the model could extract more relevant features, including spatial information, from this limited dataset. This combined approach led to improved accuracy in heart disease prediction, demonstrating its potential for scenarios with restricted data availability.

Talukdar et al.[8] investigated the potential of Artificial Neural Networks (ANNs) for predicting cardiovascular disease (CVD) in a cohort of 1,500 patients. The data was rigorously divided into training (70%), validation (15%), and testing (15%) sets to ensure generalizability. The findings suggested that ANNs held promise for early CVD detection. Further research to address limitations could lead to more accurate and generalizable models for early detection, ultimately benefiting patient outcomes.

Taylan et al.[9] investigated the application of machine learning techniques for predicting cardiovascular disease (CVD). They compared the performance of Adaptive Neuro-Fuzzy Inference Systems (ANFIS), Support Vector Machines (SVMs), and Artificial Neural Networks (ANNs) in analyzing a rich dataset of patient information. This data included demographics, medical history, lifestyle factors, clinical measurements, and even imaging data. The quality of the dataset, featuring a variety of well-described patient data and diagnoses, paved the way for the development of efficient machine learning models for CVD prediction, potentially leading to significant advancements in preventative healthcare.

Alqahtani et al.[10] explored the potential of ensemble-based machine learning methods for accurate cardiovascular disease (CVD) prediction. Their study focused on leveraging techniques like multi-layer perceptrons, Random Forest, and K-nearest neighbors. To optimize the model's effectiveness, they identified the most relevant features from a sizable Kaggle dataset containing information on over 70,000 patients and 13 CVD-related variables. Their findings suggested that ensemble-based machine learning approaches held promise for developing precise and effective CVD prediction models. With an accuracy of 88.70%, the suggested ensemble learning model, however, surpassed all other techniques.

Swathy et al.[11] investigated the use of Support Vector Machines (SVMs) and Artificial Neural Networks (ANNs) for detecting early signs of cardiovascular disease (CVD). The research leveraged data from a well-established repository, likely the UCI Cleveland Heart Disease Database, to train and evaluate their models.

IARJSET

ISSN (Online) 2393-8021
ISSN (Print) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

3rd-International Conference on Muti-Disciplinary Application & Research Technologies (ICMART-2024)

Geetanjali Institute of Technical Studies

Vol. 11, Special Issue 2, May 2024

Their work highlighted the potential of machine learning approaches like SVMs and ANNs to create more precise and personalized CVD risk assessments, paving the way for improved preventative healthcare strategies.

Indrakumari et al.[12] highlighted the importance of accurate heart disease prediction using machine learning and big data analysis. Their research explored K-means clustering techniques to identify recognized risk factors for the disease. The study emphasized the potential of combining machine learning with big data from various research datasets to improve patient outcomes and the diagnosis of cardiac disease.

Pasha et al.[13] investigated the use of machine learning for cardiac disease prediction. They compared the performance of SVM, KNN, Decision Trees, and ANN on a cardiac attack dataset. By focusing on improving model accuracy and reliability, they achieved an accuracy of 85.24% on a dataset of cardiac attacks. This study demonstrated the potential of machine learning for effectively predicting cardiac disease.

Table 1.1 given below summarizes the researchers' work on Cardiovascular Disease Prediction.

| **Researcher Name + Year** | **Model used** | **Purpose** | **Dataset** | **Result** |
|---|---|---|---|---|
| Rahman et al.2024 | Self-attention-based transformer model, hyOPTXg (improved XGBoost) | To investigate how deep learning (DL) and machine learning (ML) can improve the prediction of cardiovascular disease (CVD). | Cleveland dataset from UCI Machine Learning repository | **Results:** The self-attention-based transformer model achieved the highest accuracy (96.51%) on the Cleveland dataset. Other models performed as follows:<br>• hyOPTXg (improved XGBoost): 94.7% accuracy<br>• XGBoost: 80.6% accuracy<br>• K-Nearest Neighbors (KNN): 90.78% accuracy |
| Mansoor et al. 2023 | Long Short-Term Memory (LSTM) | The research explored the use of deep learning models, specifically LSTMs | • Cleveland dataset from the UCI Machine Learning Repository. | LSTMs achieved the highest accuracy (91%) compared to:<br>• Artificial Neural Networks (ANNs) - 90% accuracy<br>• Convolutional Neural Networks (CNNs) - 88% accuracy |
| Dhaka et al. 2023 | The model is a hybrid deep learning model | The research aims to address security and efficiency limitations in existing disease prediction models | The study uses two public datasets | • The proposed model achieves high accuracy (above 97%) on both public datasets.<br>• It outperforms previous models in terms of security and potentially other metrics |
| Demir et al. 2023 | • Convolutional Neural Networks (CNNs)<br>Deep Neural Networks (DNNs) | The research aims to explore the application of deep learning (DNNs and CNNs) for heart | A dataset from Kaggle containing various features related to heart disease. | Both CNN and DNN models achieve good accuracy (over 93%), outperforming some prior research in terms |

| | | | | |
|---|---|---|---|---|
| | | disease prediction, addressing a gap in previous studies that mainly focused on machine learning techniques. It emphasizes the importance of early detection for this critical health issue. | | of accuracy metrics |
| Sharma et al. 2023 | The research proposes a deep learning model for coronary heart disease (CHD) prediction | The research aims to develop a deep learning model for CHD prediction that can be applied in real-world hospital settings. It emphasizes improving accuracy and reducing false negatives compared to traditional machine learning methods. | A large US survey dataset, BRFSS-2015 (Behavioral Risk Factor Surveillance System - 2015). | The proposed model achieves high accuracy (98.28%). |
| Gopalakrishnan et al. 2023 | Convolutional Neural Network (CNN) | The research highlights the importance of early heart disease detection and proposes a CNN model as a valuable tool for this purpose. CNNs are effective in extracting relevant information from raw medical data for disease prediction | The dataset of patients with heart illness is used to assess the efficacy of the suggested CNN technique. | The study mentions improved accuracy for the CNN model |
| García-Ordás et al. 2023 | Sparse Autoencoder ( SAE) and Convolutional Neural Network(CNN) | Achieving high accuracy in heart disease prediction using datasets with limited features. | Includes 918 samples with 11 clinical features per patient. | The joint training of SAE and CNN allows the model to extract more relevant features with spatial location information, leading to improved accuracy. |
| Talukdar et al. 2023 | Artificial neural networks (ANNs) | to forecast the risk of cardiovascular disease (CVD) in 1,500 patients | The dataset of 1,500 cardiac patients is divided into three sets: training data (70% of the total), validation data (15%), and testing data (15%). | This study suggests that ANNs may be useful for early CVD prediction. Resolving the issues may result in early detection models that are more precise and broadly applicable, which will eventually be advantageous to patients. |
| Taylan et al. 2023 | Adaptive Neuro-Fuzzy Inference Systems (ANFIS), Support Vector Machines (SVMs), and Artificial | To predict cardiovascular disease (CVD) | Patient demographics (age, ethnicity) Medical history (prior conditions, medications) | Excellent datasets with a variety of patient data and precisely described diagnoses for efficient ML model training in |

  

| | | | | |
|---|---|---|---|---|
| | Neural Networks (ANNs) | | Lifestyle factors (smoking, diet, exercise) Clinical measurements (blood pressure, cholesterol, glucose) Imaging data (ECG, echocardiograms | CVD prediction studies |
| Alqahtani et al. | multi-layer perceptrons, Random Forest, and K-nearest neighbor. | most pertinent features for prediction to get the data ready for analysis | a sizable Kaggle dataset that includes details on more than 70,000 patients and 13 CVD-related variable | Ensemble-based machine learning techniques, especially the one put out in the study, have the potential for precise and effective CVD prediction |
| Swathy et al. 2022 | Support vector machines and artificial neural networks | The detection of early CVD signs | UCI repository or the Cleveland Heart Disease Database | researchers may create more precise and individualised CVD risk assessments |
| Indrakumari et al. 2020 | K-means clustering | lists recognized risk factors and emphasize the significance of accurately predicting heart disease | Various research datasets | machine learning combined with big data might enhance patient outcomes and diagnosis of cardiac disease |
| Pasha et al. 2020 | SVM, KNN, Decision Trees, and ANN | Improving the models' accuracy and dependability | 85.24% on a dataset of cardiac attacks | Demonstrating the potential of machine learning for the prediction of cardiac disease. |

Table 1.1 Researchers Work on Cardiovascular Disease Prediction

## III. CONCLUSION

It becomes clear from reviewing previous studies that there are far too many benefits of machine learning and deep learning that can be used in the healthcare industry. Results from a clever and comprehensive learning process can be on par with those of medical professionals.

Health workers can more easily forecast sickness at an early stage by utilizing the chance to uncover hidden patterns and linkages that can only be found by utilizing already-existing data. Currently, available options for machine learning and deep learning can be used as an "elevator" to assist in the development of cutting-edge systems. These options offer a variety of approaches and modes of operation, allowing researchers to select their preferred algorithms and combine them with other techniques to maximize the performance of the model.

The bulk of the study's findings demonstrated that deep learning could even outperform machine learning in terms of output quality. Large volumes of datasets can be used simultaneously with its support, boosting the amount of data that can be learned at once to yield excellent results. The deep learning model routinely produces high-quality results, with 85% and above. Nevertheless, it is also thought that the development of combination approaches on a single model can yield accurate forecasting outcomes. For instance, machine learning models using feature selection and other supplementary techniques can reach up to 97% maximum accuracy. Therefore, depending on the applicability and goals of the model construction, both of these methods can be used, even if the outcomes of deep learning models are superior to those of machine learning.

## REFERENCES

[1] Rahman, A. U., Alsenani, Y., Zafar, A., Ullah, K., Rabie, K. M., & Shongwe, T. (2024). Enhancing heart disease prediction using a self-attention-based transformer model. *Scientific Reports, 14*(1). https://doi.org/10.1038/s41598-024-51184-7

[2] Mansoor, C., Chettri, S. K., & Naleer, H. (2023). Efficient prediction model for cardiovascular disease using deep learning techniques. In South Eastern University of Sri Lanka, Assam Don Bosco University, & South Eastern University of Sri Lanka, *Migration Letters* (pp. 449–459) [Journal-article]. https://www.migrationletters.com

[3] Dhaka, P., Sehrawat, R., & Bhutani, P. (2023). An Innovative Approach to Cardiovascular Disease Prediction: a hybrid deep learning model. *Engineering, Technology and Applied Science Research/Engineering, Technology and Applied Science Research*, *13*(6), 12396–12403. https://doi.org/10.48084/etasr.6503

[4] Demir, E. E., Bozkurt, F., & Ayık, Y. Z. (2023). Early-stage heart failure disease prediction with deep learning approach. Journal of Scientific reports-A : (Online), 055, 34–49. https://doi.org/10.59313/jsr-a.1341663

[5] Sharma, N., Malviya, L., Jadhav, A., & Lalwani, P. (2023). A hybrid deep neural net learning model for predicting Coronary Heart Disease using Randomized Search Cross-Validation Optimization. *Decision Analytics Journal*, *9*, 100331. https://doi.org/10.1016/j.dajour.2023.100331

[6] Gopalakrishnan, S., 1, Sahaya Sheela, M., Saranya, K., & Jasmine Hephzipah, J. (2023). A novel Deep Learning-Based Heart Disease Prediction System using Convolutional Neural Networks (CNN) algorithm. *International Journal of Intelligent Systems and Applications in Engineering*, *516–522*, 516–522.

[7] García-Ordás, M. T., Bayón-Gutiérrez, M., Benavides, C., Aveleira-Mata, J., & Benítez-Andrades, J. A. (2023). Heart disease risk prediction using deep learning techniques with feature augmentation. *Multimedia Tools and Applications*, *82*(20), 31759–31773. https://doi.org/10.1007/s11042-023-14817-z

[8] Talukdar, J., & Singh, T. P. (2023). Early prediction of cardiovascular disease using artificial neural network. *Paladyn (Online)*, *14*(1). https://doi.org/10.1515/pjbr-2022-0107

[9] Taylan, O., Alkabaa, A. S., Alqabbaa, H. S., Pamukçu, E., & Leiva, V. (2023). Early Prediction in Classification of Cardiovascular Diseases with Machine Learning, Neuro-Fuzzy and Statistical Methods. *Biology (Basel)*, *12*(1), 117. https://doi.org/10.3390/biology12010117

[10] Alqahtani, A., Alsubai, S., Sha, M., Vilčeková, L., & Javed, T. (2022). Cardiovascular Disease Detection using Ensemble Learning. *Computational Intelligence and Neuroscience*, *2022*, 1–9. https://doi.org/10.1155/2022/5267498

[11]Swathy, M., & Saruladha, K. (2022). A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques. ICT Express, 8(1), 109–116**. https://doi.org/10.1016/j.icte.2021.08.021**

[12] Indrakumari, R., Poongodi, T., & Jena, S. R. (2020). Heart Disease Prediction using Exploratory Data Analysis. Procedia Computer Science, 173, 130–139. **https://doi.org/10.1016/j.procs.2020.06.017**

[13] Pasha, S. N., Ramesh, D., Mohmmad, S., Harshavardhan, A., & Shabana, N. (2020). Cardiovascular disease prediction using deep learning techniques. IOP Conference Series. Materials Science and Engineering, 981(2), 022006. **https://doi.org/10.1088/1757-899x/981/2/022006**