International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 OPEN ACCESS
Available Online at : www.ijsrcseit.com
doi : https://doi.org/10.32628/IJSRCSEIT

Heart Disease Prediction and Diagnosis Using Deep Learning Techniques

Bhaludra R Nadh Singh
Professor of CSE & Head, Department of Computer Science and Engineering, Bhoj Reddy Engineering College for Women, Vinay Nagar, Hyderabad. Telangana, India

ARTICLEINFO

Article History:

Accepted: 02 Jan 2024
Published: 22 Jan 2024

Publication Issue
Volume 10, Issue 1
January-February-2024

Page Number
251-257

ABSTRACT

Heart disease is still a major global health concern, but its effects can be significantly diminished via early detection and prevention. Using a combination of established risk variables and cutting-edge machine learning methods and For the predicted accuracy, a machine learning ensemble model is used. This ensemble model combines neural networks, support vector machines, and decision trees to accurately represent intricate nonlinear interactions between the variables. Validation of sizable and varied patient dataset, including both those with and without cardiac disease, is used to run the model. The model's performance indicators, such as accuracy, precision, recall, and F1-score, are used to assess how well it can identify those who are at high risk. To demonstrate the improvements in accuracy and dependability, comparisons are done against current risk prediction models. The findings show the integrated model's improved predictive power and offer clinicians and other healthcare professionals a useful tool for identifying those who are more likely to acquire heart disease. By providing insights into the main causes of the disease, this method may help develop customized preventive measures. The suggested integrated predictive model has promise for enhancing public health outcomes and fostering proactive cardiac treatment because heart disease continues to be a major cause of mortality.

Keywords : IoT, Machine Learning, Decision Tree, PPG Sensors, SVM, Model Prediction

I. INTRODUCTION

Heart disease is a major global health concern, and combining machine learning (ML) and internet of things (IoT) technology can improve early identification and prevention. Healthcare systems can use real-time data gathering capabilities from IoT-enabled devices to gain a more accurate picture of a

Copyright © 2023 The Author(s): This is an open-access article distributed under the terms of the Creative Commons Attribution **4.0 International License (CC BY-NC 4.0)** which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

person's cardiovascular health. ML algorithms can adapt to a patient's changing condition, enabling early detection of anomalies and trends that could indicate heart disease onset.

Remote monitoring through IoT devices (IOT devices like smart-watch that have integrated heart rate monitoring sensors like PPG (Photoplethysmography) sensors, this system forecasts the likelihood of developing heart disease. In the traditional risk assessment, factors including age, gender, blood pressure, cholesterol, and smoking history are included. This reduces the need for in-person medical visits, enhancing patient convenience and enabling healthcare professionals to monitor a larger number of patients effectively.

Challenges such as data security, interoperability of IoT devices, and model interpretability need to be addressed for widespread adoption of ML-powered IoT solutions for heart disease prediction. The development and application of a heart disease prediction system that uses the Random Forest algorithm and boasts a remarkable accuracy rate of 95.2% is a significant advance in both healthcare and data science. Heart disease is still a major global health concern, highlighting the crucial role that accurate prediction models play in enabling early diagnosis and prompt care. Achieving such a high level of accuracy is an outstanding accomplishment with significant benefits for improving patient care and maximizing health outcomes.

II.RELATED WORK

In order to automate the detection of cardiac diseases, the Ensemble Deep Learning System for Health (EDL-SHS) was developed [18] in a cloud computing environment driven by incorporated Internet-driven things. In this situation, the concept of "Health Fog" provides medical care as a "fog" by utilising IoT devices to efficiently manage patient heart data in response to user requests. The fog-busting nucleus analyses the model's effectiveness by assessing variables like latency, the black band, energy consumption, error, precision, and execution time. Because of its adaptability, Health-Fog may be tailored to meet the needs of the customer and provide the greatest service quality or prediction accuracy in a variety of fog computing scenarios. In order to bridge the resource gap for the high-precision demands of deep learning, complex deep learning networks are easily incorporated into edge computing paradigms using special communication mechanisms and models, such as assembly. This ensures increasing precision with minimal delays.

To quickly identify heart disease, recurrent networks of neurons (RNN) were introduced. Their new neural network models quickly recognise events over the course of 20 to 18 months of monitoring by include cases and controls. Using neural networks, vectorial support systems, and a K-nearest neighbour classifier, the model's performance indicators were compared to those of a regularised linear regression. The emphasis of the design is on using temporal relationships in conjunction with deep learning models, particularly over a finite period of observation of between 12 and 18 months. The ability to avoid unanticipated cardiac accidents is improved as a result. Il has introduced a Deep Neural Network (DNN) focus for diagnosing cardiac diseases [20]. Their research yielded significant discoveries while displaying a DNN architecture with five levels that was created to minimise and optimise algorithmic risk. Additionally, the architecture based on optimisation manages data flaws and errors effectively while delivering outstanding performance. The research's optimised structures were assessed during the evaluation phase using a K-vertex cross-validation and a Matthews correlation coefficient (MCC) evaluation. Through the use of open-source software and a publicly accessible data base from the Cleveland Clinic, the study demonstrated the use of DNN in the medical field. Utilising an adjustable system built on vague rules is a new way to assess the threat posed by the cardaca disease. The automatic diagnosis system makes use of a genetic algorithm and an improved particulate variable optimisation technique with an exceptional 92.3% accuracy level [21]. By combining methods for selecting multiple and univariate characteristics with a decision tree for classification, an additional technique for identifying cardiac diseases achieves a high level of accuracy of 92.8% [22].Furthermore, a sequential forward selection (SFS) feature selection method and a random forest classification method are combined in an IoT-driven hybrid system for the prediction of cardiovascular illness. With a remarkable 98% accuracy compared to other heuristic recommender systems, this holistic method not only offers precise forecasts but also offers age-and gender-specific physical and dietary advice [27]. The Kernel random forest [28], a data-driven ensemble classifier, has extraordinary performance by obtaining 98% accuracy on a heart disease dataset.A ground-breaking framework for the Internet of Things that uses deep convolutional neural networks and wearable sensors to collect blood pressure and ECG data outperforms logistic regression and other neural networks with a higher accuracy of 98.2% [29]. The chance of getting heart disease is also predicted by a sophisticated smart system that examines information from wearable sensors and patient medical histories. It is possible to identify heart disease with an amazing 98.5% accuracy using a feature fusion technique and the ensemble deep learning model logistics, and you can also get personalised eating advice based on your medical problems [30].

III.PROPOSED SYSTEM

DATASET:

The UC Irvine ML Repository-Cleveland dataset, which contains 303 instances and 14 attributes, is included in this dataset.

Features of the dataset:

• Age: Patient age.

• Sex: For males indicates 1, females 0.

• Chest pain type: For angina - (0), atypical angina – (1), non-anginal pain – (2), asymptomatic –(3).

• Resting blood pressure: Resting blood pressure upon hospital admission. Measured in mm/Hg.

• Serum Cholesterol: Blood cholesterol level measured in mg/dL.

• Fasting blood sugar: If the blood sugar level is over 120 mg/dL after a fast of not eating overnight, it is considered to be high (1—true). If it is below 120 mg/dL, it is considered to be normal (0—false).

• Resting ECG: An ECG test result can be categorized as follows: 0 for a normal result, 1 for the presence of ST-T wave abnormality, and 2 for left ventricular hypertrophy.

• Maximum heart rate: Maximum heart rate during exercise.

• Exercise angina: Angina occurred by a workout, 0 for no; 1 for yes

• Old peak: ST depression due to exercise relative to relaxation will observe in the ECG test.

• ST slope: Maximum workout 1 for upsloping; 2 for flat; 3 for down sloping

• Ca: The number of major blood vessels that can be visualized using fluoroscopy can range from 0 to 3.

• Thal: Thalassemia is a blood disorder caused by abnormal hemoglobin production, with a score of 3 indicating normal production, 6 indicating permanent deficiency, and 7 signifying temporary impairment.

• Target: No heart disease—0; heart disease—1

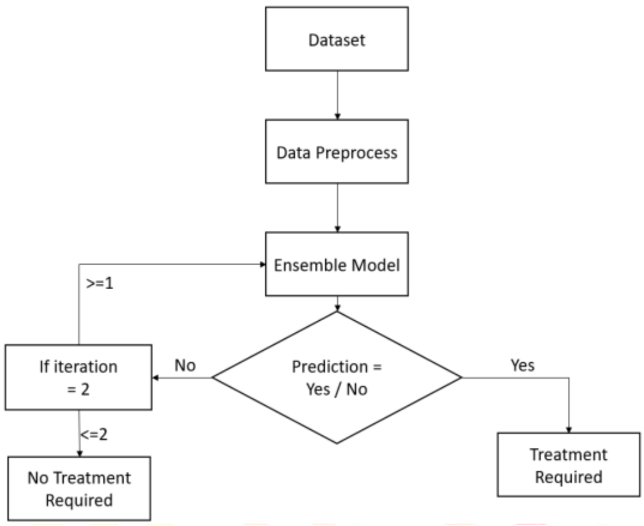

Fig 1: workflow of heart prediction system

The workflow of the prediction system in the image can be summarized as follows:

1. The user enters the desired data. Data about patients who have heart disease is collected from medical records, surveys, or other sources.
2. The data is pre-processed to remove noise, outliers and missing values.
3. The data is split into training and testing sets.
4. A machine learning algorithm is trained on the training set.
5. The trained model is used to make predictions on the testing set.
6. The accuracy of the model is evaluated.
7. The model is deployed and used to make predictions on new data.

The Fig.1 shows a specific example of a prediction system that is used to predict the treatment required for a patient. The system takes as input the patient's medical records, and then uses an ensemble model to make a prediction about the best treatment. The ensemble model is a combination of multiple machine learning algorithms, which helps to improve the accuracy of the predictions.

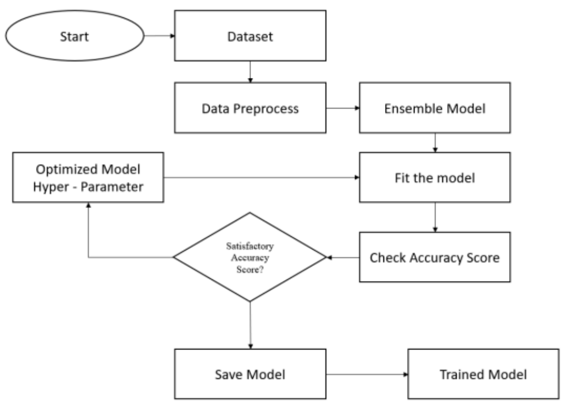

Fig 2 : Proposed Architecture

Predictive Model Uncertainty Estimation:

- Implementing methods to calculate the degree of uncertainty in model forecasts, giving users a confidence score for every forecast.

Classification Methods:

1. Random Forest Classifier:

By averaging the forecasts of their real trees, the RF classifier generates predictions. RF is a supervised machine learning technique built on ensemble learning. It combines numerous decision trees using bagging to increase prediction accuracy. Every person receives individualized bagging instruction. Each decision tree is tested during the training process using various data samples that were produced at random using replacements from the original dataset. A random selection of features is also made when building trees. The combined projections from different trees are combined by a majority vote.

Model Accuracy: 95.2%

2. KNN Classifier:

KNN is a method of lazy learning or instance-based learning. The method of developing a model without the need for training data is referred to as lazy learning. they are chosen from a collection of objects with predetermined traits or classifications. It first determines the k data points that are most similar to the new data point in order to forecast the label of the new data point. The user must choose the k value as a hyper parameter. while the k number is larger, more data points are considered while creating predictions, which could lead to slower performance but also more precise forecasts. As the k value drops, fewer data points are considered, which might lead to faster processing but less accurate projections.

Model accuracy: 84%

3. Decision Tree:

It is a supervised machine learning approach known as a decision tree can be applied to both classification and regression applications. It functions by building a decision tree-like structure, with each decision node standing in for a data feature and each leaf node for a class label. The feature that best divides the data into two pure subsets is chosen for the splitting process. A straightforward and understandable algorithm, decision trees are simple to comprehend and use. They are also incredibly adaptable and useful for a wide range of jobs. They may, however, be susceptible to data noise and overfit the training set

IV.CONCLUSION

In conclusion, this research presents a promising strategy has been the combination of machine learning (ML) algorithms and internet of things (IoT) technology. With the help of IoT data and three wellknown ML algorithms, Decision Trees, K-Nearest Neighbors (KNN), and Random Forest, this study aimed to improve heart disease prediction. The study's main accomplishment was a remarkable 95.2% accuracy rate using the Random Forest method. It was used as the main machine learning algorithm, prediction accuracy increased significantly when compared to the other algorithms. This demonstrates how well ensemble learning techniques work with challenging medical datasets. It also offers information on feature importance, which can help medical professionals comprehend the elements that have the greatest impact on heart disease. Future investigations and clinical judgments can be informed by this information. IoT and ML integration, together with the Random Forest algorithm's 95.2% accuracy rate, marks a substantial leap in the prediction of cardiac disease. By enabling early detection and intervention, this technology has the potential to improve healthcare and ultimately save lives. However, prior to the system's widespread implementation in clinical practice, it is essential to address issues and guarantee its dependability and security.

II. REFERENCES

[1]. A Angel Nancy, Dakshanamoorthy Ravindran, P M Durai Raj Vincent, Kathiravan Srinivasan and Daniel Gutierrez Reina, IoT-Cloud-Based Smart Healthcare Monitoring System for Heart Disease Prediction via Deep Learning (2022) EISSN 2079-9292, Published by MDPI

[2]. Raniya R. Sarra, Ahmed M. Dinar, Mazin Abed Mohammed and Karrar Hameed Abdulkareem, Enhanced Heart Disease Prediction Based on Machine Learning and χ2 Statistical Optimal Feature Selection Model (2022) EISSN 2411-9660, Published by MDPI

[3]. Ahmad F. Subahi, Osamah Ibrahim Khalaf, Youseef Alotaibi, Rajesh Natarajan, Natesh Mahadev and Timmarasu Ramesh, Modified Self-Adaptive Bayesian Algorithm for Smart Heart Disease Prediction in IoT System (2022) EISSN 2071-1050, Published by MDPI

[4]. Jafar Abdollahi, Babak Nouri-Moghaddam, A hybrid method for heart disease diagnosis utilizing feature selection-based ensemble classifier model generation (2022) Iran J CS Sci 5, 229–246, Published by Springer

[5]. Ahamed, Jameel; Manan Koli, Abdul; Ahmad, Khaleel; Alam Jamal, Mohd.; Gupta, B. B., CDPS-IoT: Cardiovascular DiseasePrediction System Based on IoT using Machine Learning (2022) Vol. 7, Published by Reunir

[6]. Md. Manjurul Ahsan, Zahed Siddique, Machine learning-based heart disease diagnosis: A systematic literature review (2022), Vol. 128 Article 102289 Published by Elsevier

[7]. Pierre Claver Bizimana, Zuping Zhang, Muhammad Asim and Ahmed A. Abd El- Latif, An Effective Machine Learning-Based Model for an Early Heart Disease Prediction (2023) Article 3531420 Published by Hindawi

[8]. Nadikatla Chandrasekhar and Samineni Peddakrishna, Enhancing Heart Disease Prediction Accuracy through Machine Learning Techniques and Optimization (2023), Vol. 11, Published by MDPI

[9]. Muhammed Golec, Sukhpal Singh Gill, Ajith Kumar Parlikad and Steve Uhlig HealthFaaS: AI based Smart Healthcare System for Heart Patients using Serverless Computing (2023), Published by IEEE

[10]. Federico Del Giorgio Solfa, Fernando Rogelio Simonato, Big Data Analytics in Healthcare: Exploring the Role of Machine Learning in Predicting Patient Outcomes and Improving Healthcare Delivery (2023), Published by IJCIM

[11]. Nouf Abdullah Almujally , Turki Aljrees, Oumaima Saidani, Muhammad Umer, Zaid Bin Faheem, Nihal Abuzinadah, Khaled Alnowaiser and Imran Ashraf, Monitoring Acute Heart Failure Patients Using Internet-of-Things-Based Smart Monitoring System (2023) Published by MDPI

[12]. Pronab Ghosh, Sami Azam, Mirjam Jonkman, Asif Karim, F. M. Javed Mehedi Shamrat, Eva Ignatious, Shahna Shultana, Abhijith Reddy Beeravolu and Friso De Boer, Efficient Prediction of Cardiovascular Disease Using Machine Learning Algorithms With Relief and LASSO Feature Selection Techniques (2021) Published by IEEE © 2023 IJNRD | Volume 8, Issue 10 October 2023 | ISSN: 2456-4184 | IJNRD.ORG IJNRD2310374 International Journal of Novel Research and Development (www.ijnrd.org) d564

[13]. Bzhar Ghafour Mohammed, Dler Salih Hasan, Smart Healthcare Monitoring System Using IoT, (2023), Published by ResearchGate

[14]. Sangeethalakshmi K., Preethi Angel S., Preethi U., Pavithra S., Shanmuga Priya V., Patient health monitoring system using IoT (2021), Published by Elsevier

[15]. Abderahman Rejeb, Karim Rejeb, Horst Treiblmaier, Andrea Appolloni, Salem Alghamdi, Yaser Alhasawi , Mohammad Iranmanesh, The Internet of Things (IoT) in healthcare: Taking stock and moving forward (2023), Published by ElsevieR

[16]. Suresh, S. Big Data and Predictive Analytics.Pediatr. Clin. N. Am.2016,63, 357–366.

[17]. Simpao, A.F.; Ahumada, L.M.; Gálvez, J.A.; Rehman, M.A. A Review of Analytics and Clinical Informatics in Health Care.J. Med. Syst.2014,38, 45.

[18]. Miotto, R.; Wang, F.; Wang, S.; Jiang, X.; Dudley, J.T. Deep Learning for Healthcare: Review, Opportunities and Challenges.Brief. Bioinform.2018,19, 1236–1246.

[19]. Pandey, S.; Janghel, R. Recent Deep Learning Techniques, Challenges and Its Applications for Medical Healthcare System: A Review.Neural Process. Lett.2019,50, 1907–1935.

[20]. S. Tuli, N. Basumatary, S. S. Gill, M. Kahani, R. C. Arya, G. S. Wander, and R. Buyya, "HealthFog: An ensemble deep learning based smart healthcare system for automatic diagnosis of heart diseases in integrated IoT and fog computing environments," Future Gener. Comput. Syst., vol. 104, pp. 187–200, Mar. 2020.

[21]. E. Choi, A. Schuetz, W. F. Stewart, and J. Sun, "Using recurrent neural network models for early detection of heart failure onset," J. Amer. Med. Informat. Assoc., vol. 24, no. 2, pp. 361–370, Mar. 2017

[22]. S. M. Awan, M. U. Riaz, and A. G. Khan, "Prediction of heart disease using artificial neural networks," VFAST Trans. Softw. Eng., vol. 13, no. 3, pp. 102–112, 2018.

[23]. Ahmed, H.; Younis, E.M.G.; Hendawi, A.; Ali, A.A. Heart Disease Identification from Patients' Social Posts, Machine Learning Solution on Spark. Future Gener. Comput. Syst. 2020, 111, 714–722.

[24]. Kishore, A.H.N.; Jayanthi, V.E. Neuro-Fuzzy Based Medical Decision Support System for Coronary Artery Disease Diagnosis and Risk Level Prediction. J. Comput. Theor. Nanosci. 2018, 15, 1027–1037.

[25]. Dileep, P.; Rao, K.N.; Bodapati, P.; Gokuruboyina, S.; Peddi, R.; Grover, A.; Sheetal, A. An Automatic Heart Disease Prediction Using Cluster-Based Bi-Directional LSTM (C-BiLSTM) Algorithm. Neural Comput. Appl. 2022, 1–14.

[26]. Van Pham, H.; Son, L.H.; Tuan, L.M. A Proposal of Expert System Using Deep Learning Neural Networks and Fuzzy Rules for Diagnosing Heart Disease. In Frontiers in Intelligent Computing:

Theory and Applications; Springer: Singapore, 2020; pp. 189–198.

[27]. Mehmood, A.; Iqbal, M.; Mehmood, Z.; Irtaza, A.; Nawaz, M.; Nazir, T.; Masood, M. Prediction of Heart Disease Using Deep Convolutional Neural Networks. Arab. J. Sci. Eng. 2021, 46, 3409–3422.

[28]. Jabeen, F.; Maqsood, M.; Ghazanfar, M.A.; Aadil, F.; Khan, S.; Khan, M.F.; Mehmood, I. An IoT Based Efficient Hybrid Recommender System for Cardiovascular Disease. Peer PeerNetw. Appl. 2019, 12, 1263–1276.

[29]. Muzammal, M.; Talat, R.; Sodhro, A.H.; Pirbhulal, S. A Multi-Sensor Data Fusion Enabled Ensemble Approach for Medical Data from Body Sensor Networks. Inf. Fusion 2020, 53, 155–164.

[30]. Khan, M.A. An IoT Framework for Heart Disease Prediction Based on MDCNN Classifier. IEEE Access 2020, 8, 34717–34727.

[31]. Ali, F.; El-Sappagh, S.; Islam, S.M.R.; Kwak, D.; Ali, A.; Imran, M.; Kwak, K.-S. A Smart Healthcare Monitoring System for Heart Disease Prediction Based on Ensemble Deep Learning and Feature Fusion. Inf. Fusion 2020, 63, 208–222.

[32]. Chavhan, P. G., Patil, R. V. ., & Mahalle, P. N. (2023). Context Mining with Machine Learning Approach: Understanding, Sensing, Categorizing, and Analyzing Context Parameters. International Journal on Recent and Innovation Trends in Computing and Communication, 11(4), 278–290. https://doi.org/10.17762/ijritcc.v11i4.6453

[33]. Thompson, A., Walker, A., Rodriguez, C., Silva, D., & Castro, J. Machine Learning Approaches for Sentiment Analysis in Social Media. Kuwait Journal of Machine Learning, 1(4). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/153