Proceedings of : International Conference on Emerging Trends in the field of Engineering and Technology (ICETET - 2022)
25th July, 2022 **pages: 86-90**

A MACHINE LEARNING BASED FUZZY DECISION SUPPORT SYSTEM FOR HEART DISEASE PREDICTION

#Shilpa Jala[1],Research Scholar,Dept of CSE,Email:jala.shilpa@gmail.com
#Dr.Mahesh Kandakatla[2],Asst.Professor,Dept of CSE,Email:maheshkandakatla@gmail.com
#Chaitanya Deemed to be University,Warangal,TS,INDIA.

Abstract: *In the world, heart disease is a leading cause of death, and its incidence is gradually rising. Fast economic exchange is crucial for environmental changes, the rise of unhealthy lifestyles, and the development of cardiovascular disorders. Cardiovascular Disease (CVD) is the leading cause of death as a result of limited access to healthcare services, a lack of cardiologists, and regular health screening procedures. In the current global environment, detecting heart disease in primary phase indications is a significant challenge. It can result in fatalities if not discovered in time. When cardiologists are not available in remote, semi-urban, or rural areas of developing nations, a good decision support system is crucial for early detection of heart disorders. To classify Noninvasive techniques like machine learning (ML) are reliable, accurate, and efficient for both healthy people and those with heart conditions. As a result, this paper presents FDSS for heart disease prediction based on machine learning. This system can predict heart disease quickly and accurately.*

Keywords: Heart disease, Cardiovascular Disease (CVD), Machine Learning, Fuzzy Decision Support System (FDSS).

I. INTRODUCTION

Heart and blood vessel disorders classified as cardiovascular diseases (CVDs) include coronary heart disease, CVDs, peripheral arterial disease, rheumatic heart disease, and other conditions. The leading cause of death worldwide is heart disease [1].

Different factors, such as physical inactivity, a poor diet, and increased alcohol and tobacco use, can damage the heart [2]. Adoption can help prevent heart damage by reducing the above-mentioned factors for a healthy lifestyle, such as consuming fewer salty foods, more fruits and vegetables, regular exercise, and abstaining from alcohol and tobacco.

A system of arteries and veins transports oxygen-rich blood from the heart to the rest of the body. The heart is a significant part of the human body. Fundamentally, coronary problems, high blood pressure, and sugar levels that harm the heart are the main causes of heart failure. Patients can now predict their future risk of developing heart failure thanks to advancements in medicine. It is advised that everyone, even those in perfect health, consult with a cardiologist twice a year to look for any signs of heart failure.

There are a number of warning signs, including difficulty falling asleep, an irregular heartbeat, swollen legs under certain conditions, and weight gain of up to 1-2 kg per day. It can be difficult to diagnose the problem because these symptoms, particularly in older people, are similar to those of many diseases, which can quickly result in death [3].Health facilities with all the necessaryEquipment for blood tests, chest X-rays,and magnetic resonance imaging (MRI).

Multiple-gated acquisition scanning (MUGA), Electro Cardio Gram (ECG), body analysis, physical strain assessment, cardiac catheterization radionuclide ventriculography, and other tests are performed and reveal various important data that aid physicians in the detection and understanding of patient heart failure damage. Early on in the progression of heart failure disease, a large number of neurohormonal regulatory mechanisms are induced (HFD). After pulmonary and cardiac remodeling, the compensatory process during the low period results in heart failure causes of ventricular dysfunction, exertional dyspnea, peripheral edoema, and lifelong changes in load and preload. More treatment options, lifestyle modifications, and implantable or drug systems like a pacemaker or defibrillator are offered to patients with heart disease.

It is important to ensure follow-up in this population because hospitalizations due to acute heart failure deterioration are a significant source of medical spending. According to statistics and analysis, heart disease, particularly heart failure, is the most significant issue that affects people. As with other disorders, the first stage of treatment and medical care for heart disorders is early detection and recognition.

Clinical decision support systems assist clinicians in making accurate diagnoses and prescribing effective treatments [4]. To ensure that these systems are operating safely and effectively, calculations must be made. Here, a Medical Diagnostic Support System (MDSS) for the evaluation of heart disorders must be introduced and put into place. To make better clinical decisions, medical diagnostic research must be highly accurate and efficient. Although the conventional MDSS has demonstrated its ability to address the majority of detection issues, it offers low.

It lacks the ability to make an exact detection due to accuracy issues. A recent development in artificial intelligence (AI) is the use of machine learning (ML) in therapy systems and medical diagnosis [5]. Scientific disciplines like economics, applied sciences, biology, and medicine have been governed by these techniques.

Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), Artificial Neural Network, Decision Tree, Logistic Regression (LR), AdaBoost, Naive Bayes, and rough set were implemented with different examiners and widely used for CVD detection.

As a result of the ML-based health decision model, the ratio of heart disorder morality has been decreased.

It is frequently recommended to use a prediction model for improved diagnosis and early detection of heart disorders in high-risk individuals to lower mortality rates and enhance decision-making for more sophisticated treatment and prevention. Therefore, an ML-based FDSS for heart disorder prediction is introduced in this study.

II. LITERATURE SURVEY

Yar Muhammad, Muhammad Tahir, Maqsood Hayat & Kil To Chong et. al., [6] presented an Earlier and exact identification and recognition of heart disorder using an intelligent computational design. In this examination, different ML classification algorithms were examined. To separate notconnected and noisy information from the obtained feature space, four different feature selection models are implemented and outputs of every model are examined including classifiers.

HFD, Readmission, and Death rate Prediction Using Machine Learning and Artificial Intelligence Algorithms was presented by Aixia Guo, Michael Pasque, Francis Loh, Douglas L. Mann, Philip R. O. Payne, et al. [7]. ML has been improved for HF diagnostic and result prediction using a set of Several factors obtained from EHR data, along with demographic, clinical, image, and medical note data, were combined to produce expert-comparable prediction results.

A Hybrid Intelligent System Structure for Prediction of Heart Disorder Utilizing ML Methods was presented by Amin Ul Haq, Jian Ping Li, Muhammad Hammad Memon, Shah Nazir, and Ruinan Sun, et al., [8]. By using a dataset, they improved an ML-based detection for CVD prediction. It used cross-validation, three models for selecting attributes, and seven well-known ML models.

Computer Aided Decision Producing CVD Detection using Hybrid NN was presented by Zeinab Arabasadi, Roohallah Alizadehsani, Mohamad Roshanzamir, Hossein Moosaei, Ali Asghar Yarifard, et al. [9]. The suggested method can improve the performance of a neural network (NN) by precisely 10% by creating the starting mass of the network using a genetic model that recommends the optimal weights.

A novel Tier- 3 Internet of Things (IoT) by MLmodel was presented by Priyan Malarvizhi Kumar, Usha Devi Gandhi, et al., [10] for the earlier identification of CVD. This analysis recommendsscalable three-tier designed to store and process a lot of data from wearable sensors. Tier 1 focuses on information gathering from wearable IoT sensor devices. For storing a large volume of wearable Internet of Things sensor data in the cloud, Tier-2 uses Apache HBase. Additionally, Tier-3 makes use of Apache Mahout to improve the LR-based algorithm for heart disorder prediction.

Diagnosis of CVD using GA-based trained Recurrent Fuzzy Neural Networks was presented by Kaan Uyar, Ahmet lhan et al. in [11]. (RFNN). This analysis recommends a Genetic Algorithm (GA) based trained RFNN for heart disorder detection. The likelihood of the root mean square error.

Along with accuracy, other factors like misclassification error, specificity, sensitivity, precision, and F-score are taken into account.

Using the Principal of Component Analysis, Dhomse Kanchan B, Mr. Mahale Kishor M, et al. [12] presented an analysis of ML Models for Special Disease Prediction. This implies that the goal of the test is to evaluate supervised ML techniques for heart disorder detection.

Tole Sutikno, Imam Much Ibnu Subroto, Goli Arji, Moloud Abdar, Sharareh R. Niakan Kalhori, and others, [13] Compared the use of data mining models for CVD prediction. The task at hand applies and contrasts data mining techniques to damage prediction in heart disorders. Following a characteristic analysis, five models including C5.0, a neural network, an SVM, a KNN, and an LR were implemented and validated on the samples.

Innovative Artificial NN-Based Decision Support System for CVD Detection was presented by Sameh Ghwanmeh, Adel Mohammad, Ali Al-Ibrahim, et al. [14]. A series of tests were conducted using actual medical data to determine the implementation and accuracy of the suggested solution. The distinct results show that increased classification accuracy took system implementation and exactness into account.

Support Vector Machine Based Decision Support System for CVD Classification with Integer-Coded GA to Select Critical Features was presented by Sumit Bhatia, Praveen Prakash, and G.N. Pillai, et al., [15]. This analysis uses the Cleveland heart disorder database, which has 303 cases divided into 5 groups, each with 13 identified characteristics. The results of five class-separated problems show a progression toward perfect accuracy using the right feature subset.

ML BASED HEART DISEASE PREDICTION SYSTEM

In this section, ML based FDSS for heart disease prediction is presented. The workflow of presented system.

The datasets used in this analysis were found in the UCI ML repository. The Cleveland database is particularly used. Tests were run using data on Cleveland Heart disease hat was obtained from the University of California, Irvine (UCI) repository. This dataset consists of 14 characteristics, eight of which are absolute and six of which are numerial.Date of birth, gender, reached high pulse, exercise induced angina, serum cholesterol in mg/dl,fasting glucose, resting ECG output, ST depression, slope of peak exercise ST segment, number of important vessels, type of chest and B.P.ache, and presenting CVD Condition are among the 14 features in this dataset Six numbers from the Cleveland dataset are missing. The Multivariate Imputation by Chained Expressions algorithm is used to impute those values. The imputation is carried out repeatedly by the algorithm.

This implies that the missing information is random. To assume the value of a missing attribute from the dataset's left, the LR method is used.

THE WORKFLOW OF PRESENTEDSYSTEM

FDSS is based on pre-processed datasets to get good estimation of risk level. For everyattributes, it considers "lower the dependency, higher the inconsistency". This helps enhance exactness and speed of Further processing by discarding features which are not related. The main objective of fuzzy based decision support system is receiving higher diagnostic accuracy,ignoring transparency problem. Therefore, fuzzy model is acquired by these systems is a type of black-box that is used only for evaluating instead of inference prediction label. The model of FDSS has to examine accuracy as well as transparency.

GA is initially trained using a pre-processed data set and a few chosen features. Every rule's fitness is determined by an aim operation. "Higher the fitness, higher the strength" is a factor. Every rule is given a strength rating based on how it performs in the FDSS, and the best and strongest rules are then moved on to the next stage to create the newest rules, which are formed using elitism, crossover, mutation, and selection. This process continues from step to step until the desired outcome is achieved. High stage and benefits of the created FDSSs are used as the standard conclusion in this examination.

Following that, feature selection is carried out. A preprocessing step called feature selection aims to reduce the number of features by choosing a subset of features that offer better analogothrim in accordance with established standards. With a standard scalar, all feature coefficients introduce similar values, each with an average of 0 and a standard deviation of 1. The input dataset is implemented using a common scalar method. Classifier execution varies depending on output scaling. In this dataset, 139 instances are in group 1 and 164 instances are in group 0, respectively. Using the Synthetic Minority Oversampling Technique, group balancing is accomplished.

SMOTE is a technique oversampling artificial minority class representatives. This approach helps prevent overfitting problems brought on by arbitrary oversampling. SMOTE is a statistical technique that increases the instances of a dataset following a balanced path. The primary task is developing new instances from minority cases that already exist and are supplied as input.

The five primary waves that make up an ElectroCardioGram waveform are P, Q, R, S, and T, with U appearing on occasion. The P shows successive depolarization of the right and left atria, the QRS complex shows depolarization of the ventricles, the T shows depolarization of the ventricles, and the U shows depolarization of the papillary muscles. The QRS complex, which connects the three graphic diversions found on a typical Electrocardiogram, represents a significant portion of the information from Electrocardiograms that were under investigation. To confirm the borderline for the peak value and time interval, the gradient flow computation is used. Whether the location is inside or outside.

To carry out categorization, SVM creates a hyperplane with two planes, one for samples related to one group and the other for samples related to different groups. The hyperplane is optimised to ensure the greatest distance between the two groups. Classes of data points that are closest to the hyperplane are called support vectors. Using an SVM classifier, selected features are finally categorised. Based on ECG results, the Support Vector Machine classifier distinguishes between heart disorders that exist and those that do not. Accuracy, precision, sensitivity, and detection time are calculated as the system's execution.

I.RESULT ANALYSIS

In this part the output analysis of presented ML based FDSS for heart disorder prediction is discussed. In this system support vector Machine ML classifier is used. The performance is calculated depended on the classified instances namely True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) that termed as follows:

True Positive: If a sample is predicted correctly as positive and in fact it is positive.

True Negative: If a sample is correctly predicted as negative and actually it is negative.

False positive: If a sample is not correctly compared to NB,and RFclassifiers predicted as negative but in fact it is positive.

False Negative: If a sample is predicted incorrectly as positive but actually it is positive.

Based on these values the performance metrics like sensitivity, precision, accuracy, specificity and F1-score are measured for performance evaluation of presented system.

Sensitivity: It is also called as TP Rate or Recall. It is a execution parameter that calculates capability of the system to make exact better predictions and is expressed.presented SVM classifier has greater performance in terms of accuracy, precision and sensitivity. Accuracy, Precision, Recall Comparison of NB, RF and SVM classifiers with various thresholds. X-axis shows classification and Y-axis represents percentage (%). The SVM has high accuracy compared to NB and RF.

Accuracy: It is an execution characteristic that calculates the capability of system to create exact predictions and is expressed.

The time comparison between presented, NB and RF based heartdisease prediction systems. The SVM based system takes less time for prediction of heart disorder compared to Naïve Bayes.

The performance evaluation of presented heart disease prediction system and other ML classifiers depended on heart disorder prediction systems is represented and Random forest. The Y-axis presents time in ms (milli seconds) and X-axis represents the classifiers.Therefore presented ML based FDSS for heart disease prediction has better results compared to different ML classifiers in form of accuracy, prediction time, precisionand sensitivity.

CONCLUSION

This analysis presents an ML-based FDSS for the prediction of heart disease. Here, SVM classifier is utilised for precise heart disease prediction. Cleveland Dataset on heart disease obtained from the UCI repository is used to carry out experiments. The fuzzy decision support system is used to predict heart disease and also takes the system's decision accuracy and transparency into account. The SMOTE method is employed to maintain system balance. ECG waveform is used to predict the patient's condition (PQRST). Finally, if a patient has heart disease, the heart disease is predicted to be positive; otherwise, the patient is predicted to be normal. The results show that the system is accurate and produces reliable results when compared to other systems.ML classifiers in form of accuracy and prediction time.

REFRENCES

[1] Gizeaddis Lamesgin Simegn, Worku Birhanie Gebeyehu, Mizanu Zelalem Degu, "Computer-Aided Decision Support System for Diagnosis of Heart Diseases",Research Reports in Clinical Cardiology,2022:133954,doi.org/10.2147/RRCC.S366380

[2] Umarani Nagavelli, Debabrata Samanta and Partha Chakraborty, "Machine Learning Technology-Based Heart Disease Detection Models", Hindawi, Journal of Healthcare Engineering Volume 2022, ArticleID7351061,9pages,doi.org/10.1155/2022/7351061

[3] Rohit Bharti, Aditya Khamparia, Mohammad Shabaz, Gaurav Dhiman, Sagar Pande, and Parneet Singh, "Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning", Computational Intelligence and Neuroscience,Volume 2021,ArticleID8387680,11pages,doi.org/10.1155/2021/8387680

[4] Pooja Rani, Rajneesh Kumar, Nada M.O. Sid Ahmed, Anurag Jain, "A decision support system for heart disease prediction based upon machine learning", Springer 2021, doi.org/10.1007/s40860-021-00133- 6.

[5] Oumaima Terrada, Soufiane Hamida, Bouchaib Cherradi, Abdelhadi Raihani1, Omar Bouattane, "Supervised Machine Learning Based Medical Diagnosis Support System for Prediction of Patients with Heart Disease", Advances in Science, Technology and Engineering Systems Journal Vol. 5,No.5,269-277 (2020), doi.org/10.25046/aj050533.

[6] Yar Muhammad, Muhammad Tahir, Maqsood Hayat & Kil To Chong, "Early and accurate detection and diagnosis of heart disease using intelligent computational model", Scientific Reports,(2020)10:19747, doi.org/10.1038/s41598- 020-76635-9.

[7] Aixia Guo, Michael Pasque, Francis Loh, Douglas L. Mann & Philip R. O. Payne,"Heart FailureDiagnosis,Readmission,and Mortality Prediction Using Machine Learning and Artificial Intelligence Models",Current Epidemiology Reports (2020) 7:212–219,doi.org/10.1007/s40471-020-00259-w.

[8] Amin Ul Haq , Jian Ping Li , Muhammad Hammad Memon, Shah Nazir, and RuinanSun, "A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms", Hindawi, Mobile Information Systems Volume 2018, Article ID 3860146,21 pages, doi.org/10.1155/2018/3860146.

[9] Zeinab Arabasadi, Roohallah Alizadehsani, Mohamad Roshanzamir, Hossein Moosaei, Ali Asghar Yarifard, "Computer Aided Decision Making for Heart Disease Detection using Hybrid Neural Network - Genetic Algorithm", Computer Methods and Programs in Biomedicine,doi:10.1016/j.cmpb.2017.01.004.

[10] Priyan Malarvizhi Kumar, Usha DeviGandhi, "A novel three-tier Internet of Things architecture with machine learning algorithm for early detection of heart diseases", Computers and Electrical Engineering 0 0 0 (2017) 1–14.

[11] Kaan Uyar, Ahmet İlhan, "Diagnosisof heart disease using genetic algorithm based trained recurrent fuzzy neural networks", 9th International Conference on Theory and Application of Soft Computing, Computing with Words and Perception, ICSCCW 2017, 22-23 August 2017,Budapest, Hungary.

[12] Dhomse Kanchan B, Mr. Mahale Kishor M, "Study of Machine Learning Algorithms for Special Disease Prediction using Principal of Component Analysis", 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication.

[13] Moloud Abdar, Sharareh R. Niakan Kalhori, Tole Sutikno, Imam Much Ibnu Subroto, Goli Arji, "Comparing Performance of Data Mining Algorithms in Prediction Heart Diseases", International Journal of Electrical and Computer Engineering (IJECE), Vol. 5, No. 6, December 2015, pp. 1569~1576, ISSN: 2088-8708 Sameh Ghwanmeh, Adel Mohammad, Ali _Al-Ibrahim, " Innovative Artificial Neural Networks-Based Decision Support System for Heart Diseases Diagnosis", Journal of Intelligent .

The performance evaluation of presented he Learning Systems andApplications, 2013, 5, 176-183, doi.org/10.4236/jilsa.2013.53019

[14] Sumit Bhatia, Praveen Prakash, andG.N. Pillai, "SVM Based Decision Support System for Heart Disease Classification with Integer-Coded Genetic Algorithm to Select Critical Features", Proceedings of the World Congress on Engineering and Computer Science 2008, WCECS 2008, October 22 - 24, 2008, San Francisco, USA.