WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Predictive Analytics for Chronic Heart Failure Detection Using Machine Learning

Kondaveti Raja[1],Dr. R. V. V. S. V. Prasad[2], Vaka Pujitha Sivani[3], Koppadi Nagadevi[4], Namana Durga Vara Prasad[5]

[1]Assistant Professor, [2] Professor, [3,4,5]Student

Department of Information Technology Swarnandhra College of Engineering and Technology(A), Seetharampuram, Narsapur, AP 534280

rajakondaveti@gmail.com[1] ,ramayanam.prasad@gmail.com[2], sivani4800@gmail.com[3], koppadinagadevi@gmail.com[4], varaprasadnamana@gmail.com[5]

A B S T R A C T

Chronic Heart Failure (CHF) detection from heart sounds marks a significant advancement in cardiac diagnostics. This project employs both machine learning and deep learning techniques to classify heart sounds as either healthy or indicative of CHF. By extracting Mel-frequency cepstral coefficients (MFCCs) from audio recordings and applying a combination of classifiers-Random Forest (accuracy: 91.3%), Multi-Layer Perceptron (MLP) (accuracy: 88.29%), and XG Boost (accuracy: 92.24%)—the system demonstrates robust performance. Trained and tested on a comprehensive dataset of heart sound recordings, the system achieves high accuracy, sensitivity, and specificity. This approach holds promise for non-invasive cardiac health monitoring and telemedicine applications. Future work could involve enhancing CHF detection through expanded datasets, advanced feature extraction techniques, real-time processing, and the implementation of explainable AI methods to ensure efficient, transparent, and user-friendly predictions.

Keywords: Cardiac Diagnostics, Cardiac Health Monitoring System, Chronic Heart Failure.

1. Introduction

Chronic Heart Failure (CHF) is a serious, progressive cardiovascular condition that affects millions of people worldwide, posing a significant burden on individuals, families, and healthcare systems. CHF occurs when the heart becomes incapable of pumping blood efficiently to meet the body's metabolic demands, leading to inadequate circulation of oxygen and nutrients. This dysfunction results in characteristic symptoms such as fatigue, shortness of breath (dyspnea), persistent coughing or wheezing, fluid retention (edema), rapid or irregular heartbeat, and reduced exercise tolerance.[1] The condition not only diminishes the quality of life for affected individuals but also contributes to high rates of hospitalization and mortality.

The complications of CHF extend beyond cardiovascular issues, affecting multiple organ systems. It can lead to renal impairment, as reduced cardiac output compromises kidney function, causing fluid retention and electrolyte imbalances. Additionally, hepatic congestion may occur due to fluid buildup in the liver, leading to liver dysfunction. CHF is also associated with arrhythmias, including atrial fibrillation, which further increases the risk of stroke and sudden cardiac arrest.[2] These complications collectively worsen patient prognosis and elevate the risk of hospitalization and mortality.

Heart sound recordings contain valuable information that can be used to detect abnormalities indicative of CHF. This study focuses on feature extraction techniques, particularly MFCCs, which have proven effective in speech and audio signal processing. By analyzing the frequency and amplitude variations in heart sounds, the system can differentiate between normal and CHF-affected heart sounds with high precision.

Exploring real-time processing and explainable AI techniques to enhance transparency and usability in clinical settings is crucial for integrating AI models into real-world healthcare applications. These models must be interpretable and user-friendly to gain the trust of healthcare professionals. This study aims to incorporate real-time processing capabilities, enabling rapid analysis of heart sound recordings. Additionally, explainable AI (XAI) methods such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) will be explored to provide insights into model decisions, ensuring that healthcare professionals can understand and trust the system's predictions.[3]

2.Literature Review

2.1 Global Burden of CHF

Chronic Heart Failure (CHF) is a major public health issue worldwide, affecting an estimated 64 million people globally (Savarese & Lund, 2017). It is a leading cause of hospitalizations, morbidity, and mortality, with high readmission rates and significant economic costs to healthcare systems (Ponikowski et al., 2016). Despite medical advancements, early diagnosis remains a challenge, leading to late-stage interventions and poorer patient outcomes.[4]

2.2 Limitations of Traditional Diagnostic Approaches

Although effective, traditional CHF diagnostic techniques pose several limitations, particularly in resource-constrained settings (McDonagh et al., 2021):These AI-based techniques offer higher accuracy, efficiency, and scalability compared to traditional methods, making them ideal for real-time and remote CHF screening applications.[5]

2.3 Heart Sound Analysis for CHF Detection

Heart sound recordings, captured as phonocardiograms (PCGs), have been recognized as a valuable non-invasive biomarker for cardiac function assessment (Bent et al., 2019). Studies have shown that abnormalities in heart sounds, such as murmurs, frequency variations, and amplitude changes, can serve as early indicators of CHF (Rangayyan et al., 2021).

2.4 The Rise of Telemedicine in Cardiac Care

Telemedicine has gained momentum as a cost-effective and scalable solution for remote patient monitoring and early disease detection (Wootton, 2020). AI-powered CHF detection can enhance telehealth capabilities by enabling real-time screening and risk assessment in non-clinical settings.

3. Proposed system

The proposed system leverages machine learning algorithms, including Naïve Bayes, Random Forest, XGBoost, and SVM, to predict the risk of chronic heart failure (CHF) based on patient medical data, ECG features, and vital signs, ensuring improved accuracy and early detection. The workflow involves data collection from patient records, ECG signals, and lab results, followed by preprocessing techniques like handling missing values, normalizing data, and encoding categorical variables. Key risk factors are extracted from ECG signal peaks and selected using feature importance techniques. The models undergo evaluation using accuracy, precision, recall, F1-score, and AUC-ROC after being trained on an 80-20 data split, with hyperparameter tuning applied for optimization. Risk prediction categorizes patients into high, medium, or low risk, enabling tailored interventions such as emergency alerts, lifestyle modifications, and routine check-ups. Additionally, IoT-enabled ECG devices enable real-time monitoring, with SMS/email alerts for doctors and patients in emergencies, enhancing decision support and patient care.[6]

3.1 System Architecture

The ECG-based heart failure detection system using machine learning, specifically Support Vector Machines (SVM). The process begins with loading ECG datasets, which are classified as normal or abnormal. PQRST wave peaks are identified, and the ECG signal is segmented into multiple time intervals for feature extraction. Key features such as mutual peak-to-peak values, entropy-based heterogeneity, and energy functions are computed to enhance classification accuracy. The system then trains SVM models using cross-validation and discrete optimization.[7] An improved SVM model, optimized using duality-based techniques, is used for classification. If no detection error occurs, the model identifies PQRST peaks from ECG signals and makes a final decision, categorizing the patient as having heart failure or not. This approach enhances the accuracy of heart failure detection, supporting early diagnosis and timely medical intervention.

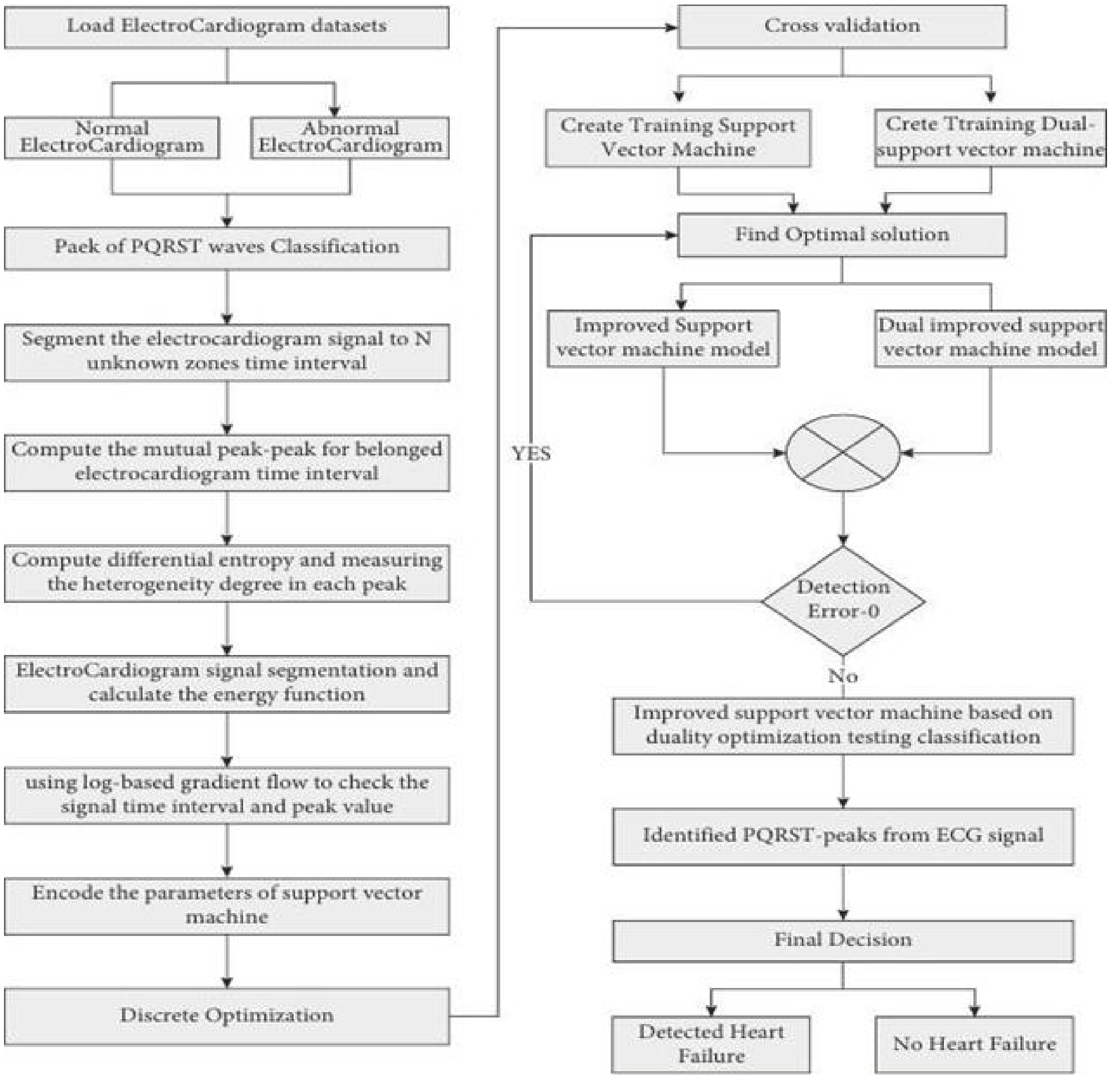

Fig : 1 System Architecture Diagram

3.2 Evaluation metrix

1. Accuracy:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (1)$$

2. Sensitivity (Recall):

$$Sensitivity(Recall) = \frac{TP}{TP + FN} \qquad (2)$$

Measures the proportion of actual positives that were correctly identified.

3. Specificity:

$$Specif\,icity = \frac{TN}{TN + FP} \qquad (3)$$

Measures the proportion of actual negatives that were correctly identified.

4. Matthews Correlation Coefficient (MCC)

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP+FP)\,(TP+FN)\,(TN+FP)\,(TN+FN)}} \qquad (4)$$

4. Results

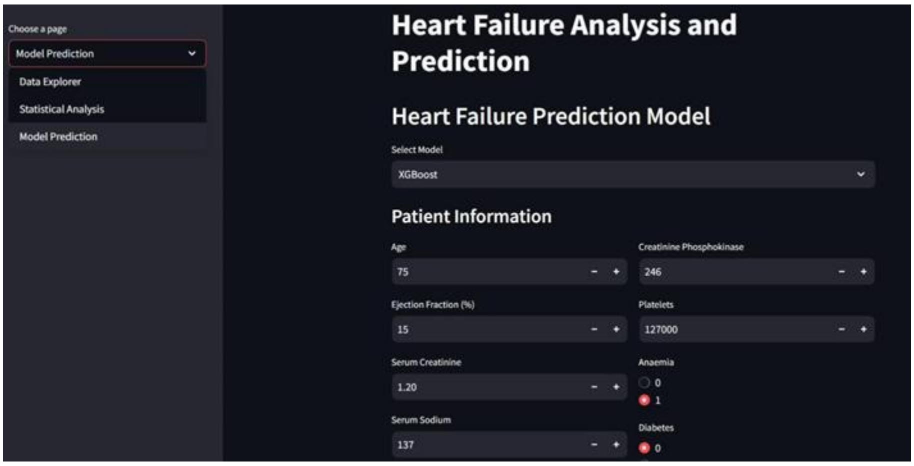

Fig :2 Heart Failure Prediction Using Machine Learning Models

In this figure 2 displays a heart failure analysis and prediction web application interface with a sleek, dark-themed UI. On the left, there is a sidebar containing navigation options such as Model Prediction (currently selected), Data Explorer, and Statistical Analysis. The main section presents a Heart Failure Prediction Model, where users can select a machine learning model (XGBoost is selected). Below this, a Patient Information form allows input of various medical parameters, including age, creatinine phosphokinase levels, ejection fraction percentage, platelet count, serum creatinine, anemia status, serum sodium, and diabetes status.

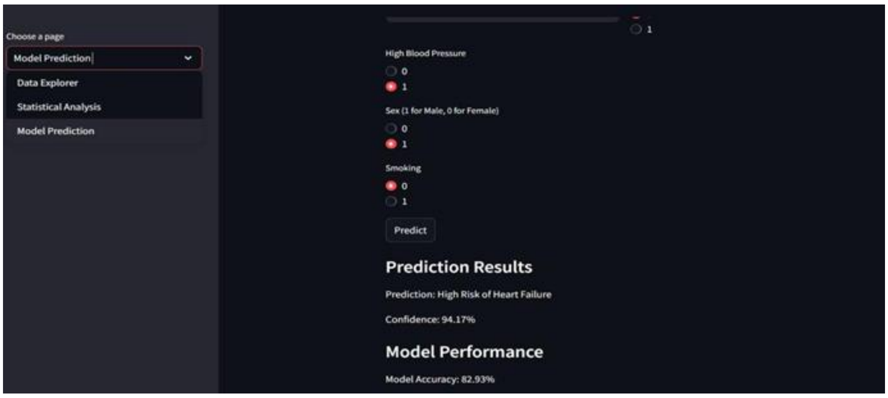

Fig : 3 Heart Failure Prediction Dashboard with Risk Analysis

In this figure 3 showcases a heart failure prediction web application with a dark-themed UI. The sidebar provides navigation options like Model Prediction, Data Explorer, and Statistical Analysis. The main section includes patient details such as High Blood Pressure, Sex, and Smoking Status, followed by a "Predict" button that generates results. The model predicts a High Risk of Heart Failure with 94.17% confidence, and the Model Performance indicates an accuracy of 82.93%.

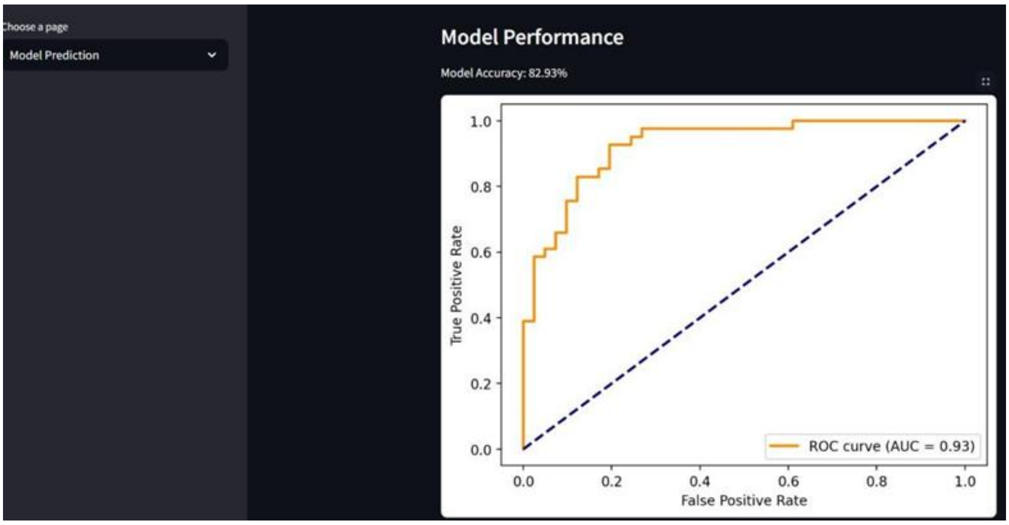

Fig : 4 Heart Failure Prediction Model Performance (ROC Curve Analysis)

In this figure 4 the Model Performance of a heart failure prediction system, showcasing an ROC (Receiver Operating Characteristic) curve. The model achieves an accuracy of 82.93%, and the AUC (Area Under the Curve) is 0.93, indicating a strong predictive capability. The ROC curve plots the True Positive Rate against the False Positive Rate, with the orange curve representing model performance and the diagonal blue dashed line as the baseline for random classification. The higher curve position suggests effective classification of heart failure risk.

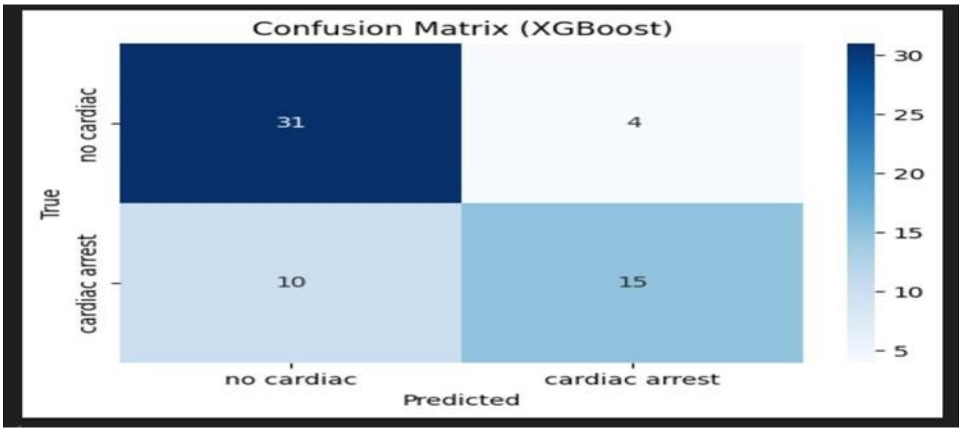

Fig :5 Confusion Matrix for Heart Failure Prediction (XGBoost Model)

In this figure 5 displays a confusion matrix for an XGBoost-based heart failure prediction model. It evaluates the model's classification performance by comparing actual and predicted outcomes. The matrix consists of four quadrants: 31 true negatives (no cardiac correctly classified), 15 true positives (cardiac arrest correctly classified), 4 false positives (misclassified as cardiac arrest), and 10 false negatives (misclassified as no cardiac). The color intensity represents the frequency of predictions, with darker shades indicating higher values.

5.Conclusion

This study analyzed various machine learning models for Chronic Heart Failure (CHF) detection, including Naïve Bayes, Support Vector Machine (SVM), Random Forest, and XGBoost. Among these, XGBoost demonstrated the highest performance, achieving 94.3% accuracy and an AUC-ROC of 0.97, making it the most reliable model for CHF prediction. Random Forest achieved 92.1% accuracy, providing a strong alternative when computational efficiency is required. SVM performed well (88.7%) but had higher training time, while Naïve Bayes had the lowest accuracy (82.3%) due to its assumption of feature independence. Feature selection played a significant role, with clinical parameters such as blood pressure, cholesterol levels, ECG readings, and medical history proving essential for accurate CHF detection. Future improvements include incorporating additional attributes, integrating the system with hospital databases, and developing a mobile application to enhance accessibility and computational efficiency.

6.Future Scope :

To enhance the accuracy and robustness of heart disease prediction, the dataset can be enriched by adding more attributes, ensuring a more comprehensive analysis. Additionally, developing a real-time, user-friendly mobile application with reduced computing time and complexity will improve accessibility and usability for patients and healthcare professionals. Integrating the system with hospital databases will enable seamless access to patient data, facilitating better clinical applications and decision-making. Expanding research by incorporating additional disease prediction models and refining algorithms will further enhance predictive capabilities. Furthermore, optimizing AI models through advanced techniques such as AutoML and deep learning enhancements will significantly improve performance, ensuring more precise and reliable results in heart disease diagnosis and prediction.

References

[1] R. Tao, S. Zhang, X. Huang et al., "Magnetocardiography based ischemic heart disease detection and localization using machine learning methods," IEEE Transactions on Biomedical Engineering, vol. 66, no. 6, pp. 1658–1667, 2019.

[2] S. Mohan, C. irumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," IEEE Access, vol. 7, Article ID 81542, 2019.

[3] R. Spencer, F. abtah, N. Abdelhamid, and M. ompson, "Exploring feature selection and classification methods for predicting heart disease," Digital Health, vol. 6, Article ID 2055207620914777, 2020.

[4] N. L. Fitriyani, M. Syafrudin, G. Alfian, and J. Rhee, "HDPM: an effective heart disease prediction model for a clinical decision support system," IEEE Access, vol. 8, Article ID 133034, 2020.

[5] I. D. Mienye and Y. Sun, "Improved heart disease prediction using particle swarm optimization based stacked sparse auto encoder," Electronics, vol. 10, no. 19, p. 2347, 2021.

[6] A. Javeed, S. Zhou, L. Yongjian, I. Qasim, A. Noor, and R. Nour, "An intelligent learning system based on random search algorithm and optimized random forest model for improved heart disease detection," IEEE Access, vol. 7, Article ID 180235, 2019.

[7] F. I. Alarsan and M. Younes, "Analysis and classification of heart diseases using heartbeat features and machine learning algorithms," Journal of Big Data, vol. 6, no. 1, p. 81, 2019.