



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

# IJRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 6, June 2024

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

[ijrset@gmail.com](mailto:ijrset@gmail.com)

[www.ijrset.com](http://www.ijrset.com)

# Application of ML in Healthcare

Samruddhi Vanaju, Shreya Sonar, Tejas V. Joshi

Department of M. C. A., Finolex Academy of Management and Technology, Ratnagiri, India

Department of M. C. A., Finolex Academy of Management and Technology, Ratnagiri, India

Assistant Professor & Head, Finolex Academy of Management and Technology, Ratnagiri, India

**ABSTRACT:** Heart disease tops the charts as the leading cause of death globally, impacting developed, developing, and underdeveloped nations alike. Early diagnosis and ongoing medical care hold the key to lowering death rates. Hence the proposed research aims to enhance Electrocardiogram (ECG) classification through a machine learning (ML) approach, addressing the imperative need for accurate detection of heart conditions [1]. Leveraging ML-libraries and Scala on the Apache Spark framework, the study tackles the challenge of irregularities inherent in ECG signals. The method achieves remarkable accuracy by extracting crucial features from ECG signals and employing various ML algorithms such as Decision Trees, Random Forests, and Gradient-Boosted Trees. The assessment of the MIT-BIH Arrhythmia and MIT-BIH Supraventricular Arrhythmia databases reveals an overall accuracy ranging from 96.75% to 98.03%. [1]. Concurrently, the research acknowledges the global burden of heart failure disease and underscores the pivotal role of ECG in early diagnosis [2]. Various machine learning techniques, including Naïve Bayes, Support Vector Machine (SVM), and XGBoost, are explored for heart disease detection and localization [2]. Additionally, an enhanced SVM method and a clinical decision support system (CDSS) incorporating density-based clustering and XGBoost are proposed to accurately predict and diagnose heart diseases [2]. This study contributes to advancing healthcare services by providing clinicians with effective tools for early diagnosis and treatment of heart conditions.

**KEYWORDS:** ECG Classification, Machine Learning (ML), Heart Disease Detection, Accuracy (or Classification Rates), Early Diagnosis

## I. INTRODUCTION

Heart disease, a silent and relentless adversary, remains a leading cause of death globally, claiming over 30% of lives each year. Early detection and intervention are critical to turn the tide in this battle. This paper proposes a multi-pronged approach that leverages the power of data analysis, cutting-edge technology, and existing healthcare infrastructure to improve patient outcomes and save lives.

The vast amount of patient data collected in electronic health records holds hidden treasures – patterns and trends that can unveil the risk of heart disease before symptoms even appear. This study proposes utilizing machine learning techniques, a form of data analysis that allows computers to "learn" from existing data. By sifting through this digital mountain range, we can identify subtle connections that might predict future heart problems. Imagine a detective meticulously examining a crime scene, piecing together seemingly insignificant details to solve the puzzle. In this case, the puzzle is the human body, and the hidden clues are indicators of potential heart trouble.

Electrocardiograms (ECGs) are a device that capture the electrical activity of the heart on the body's surface. This simple, non-invasive test is widely used in diagnosing heart disease because it offers a valuable window into the health of the cardiovascular system. Millions of people suffer from irregular heartbeats which is deadly. ECG recordings, capture the electrical activity of your heart, these contain a wealth of information about its health. This information can be broken down into different categories to get a complete picture such as ECG morphology, heartbeat interval features (temporal features), beats correlations, and summits values. To analyze and diagnose heart disease, extracting key features and classifying heartbeats are the most crucial steps.[3]

Another weapon in this fight is the decision support system (DSS). This digital sidekick for medical professionals acts as a central hub for patient data collected across various healthcare facilities. By analyzing this comprehensive data set, the DSS can provide valuable insights and suggest the most appropriate treatment plan for each individual. Imagine a doctor having access to a vast library of medical knowledge and patient history, instantly at their fingertips. This

empowers healthcare professionals to make informed decisions and personalize treatment plans for optimal patient outcomes.

Early diagnosis is paramount, especially for a condition like heart failure disease (HFD). Traditional methods like ECG scans, which provide detailed pictures of the heart's electrical activity, have long been used for diagnosing HFD. However, the world of medical technology is constantly evolving. New non-invasive diagnostic tools like MCG (Magnetocardiography) are emerging, offering even clearer and more detailed information about the heart's function. These advancements allow doctors to detect problems at an earlier stage, paving the way for more effective treatment strategies.

This study proposes a comprehensive approach to combat heart disease. By combining the power of data analysis, real-time monitoring with IoT, and advanced diagnostic tools, we can move beyond reactive treatment towards proactive prevention. Early identification of heart disease through prediction models and continuous monitoring allows for timely intervention, ultimately saving lives and improving patient care. This is a significant step towards a future where heart disease becomes less of a threat and more of a manageable condition.

## II. LITERATURE REVIEW

The human heart's electrical activity is recorded using an Electrocardiogram (ECG), a noninvasive technique that captures various waveforms through skin electrodes. This method is instrumental in identifying heart diseases as it reflects heartbeat, heart rate, and overall cardiac health. The cells within the human body rely on the cardiovascular system to facilitate their transportation needs, as they do not have direct contact with external environments.

The cardiovascular system comprises two primary fluids: blood and lymph. Blood circulates through the heart and blood vessels, forming the circulatory system. The lymphatic system, consisting of lymph nodes and lymphatic vessels, transports lymph. Together, these systems integrate to form the comprehensive cardiovascular system.

The heartbeat is orchestrated through a series of actions known as the cardiac cycle. This cycle involves the synchronized contraction of both atria, followed by the contraction of each ventricle a fraction of a second later. The heart, composed of interconnected heart muscle cells, functions cohesively—when one cell contracts, it stimulates the neighboring cells to contract as well. Between beats, the cardiac muscles relax, which is essential for aerobic respiration and maintaining the heart's functionality.

Additionally, recent studies have emphasized the importance of understanding the electrical conduction system of the heart, including the sinoatrial (SA) node, atrioventricular (AV) node, and the Purkinje fibers. These components play critical roles in regulating heart rhythms and ensuring efficient blood circulation. Advances in ECG technology have also led to improved detection and monitoring of arrhythmias, ischemic heart diseases, and other cardiac conditions, thereby enhancing diagnostic accuracy and patient outcomes<sup>[30][31]</sup>.

Furthermore, the integration of artificial intelligence and machine learning in ECG analysis is emerging as a promising avenue. These technologies enable more precise interpretation of ECG data, facilitating early detection of potential cardiac issues and providing valuable insights for preventive cardiology<sup>[32][33]</sup>.

## III. BACKGROUND

Cardiovascular diseases (CVDs), encompassing conditions such as heart attacks and strokes, remain the leading cause of death globally, accounting for over 30% of all fatalities. Projections suggest that without significant intervention, the annual death toll from CVDs could rise to 22 million by 2030. The primary factors contributing to heart disease include physical inactivity, unhealthy diet, and the consumption of alcohol and tobacco. These risk factors can be mitigated through lifestyle modifications, such as reducing salt intake, increasing the consumption of fruits and vegetables, regular physical activity, and abstaining from alcohol and tobacco use.

Recent advancements in healthcare have resulted in the generation of vast amounts of patient data, commonly referred to as Big Data, through Electronic Health Record Systems. This data provides an opportunity to apply data mining and machine learning techniques to uncover hidden patterns and develop predictive models for heart disease. Data mining is



the process of extracting useful, previously unknown information from large datasets, which is particularly valuable in the healthcare industry for disease diagnosis and prediction.

Numerous studies have demonstrated the potential of machine learning algorithms in improving the detection and prediction of heart diseases. Traditional methods, such as the models provided by the American College of Cardiology/American Heart Association (ACC/AHA), have been enhanced through the application of artificial intelligence. For instance, Zhao, Wang, and Nakahira (2011) explored the use of IoT and sensor technologies for advanced human health management, identifying key issues to be addressed. Similarly, Chiuchisan and Geman (2014) developed a home monitoring system for Parkinson's disease, demonstrating the potential for remote health monitoring and decision support systems.

Specific machine learning techniques have shown promise in heart disease detection. For example, Naïve Bayes algorithms have been employed for heart disease prediction by leveraging historical datasets like the Cleveland dataset. Additionally, Support Vector Machines (SVM) combined with XGBoost have been used to classify heart disease based on features from the frequency domain, time domain, and information theory. Improved SVM methods using duality optimization schemes have also been explored for automatic heart failure identification.

Moreover, the development of Clinical Decision Support Systems (CDSS) has been pivotal in enhancing the early diagnosis and treatment of heart disease. CDSS models incorporate advanced data processing techniques such as DBSCAN for outlier detection, SMOTE-ENN for balancing training data, and XGBoost for heart disease prediction. These systems aim to provide clinicians with tools for early detection, thereby facilitating timely intervention and improving patient outcomes.

The integration of IoT with machine learning further enhances the monitoring and management of heart disease. IoT-based systems, equipped with various physiological sensors and microcontrollers, enable real-time monitoring of cardiac health. These systems collect sensor data, analyze it, and transmit it to healthcare providers, allowing for continuous monitoring and prompt response to abnormalities.

Overall, the application of machine learning and IoT in the detection and prediction of heart disease represents a significant advancement in healthcare. These technologies offer the potential to improve diagnostic accuracy, enhance patient monitoring, and ultimately reduce the global burden of cardiovascular diseases. This research aims to develop a robust machine learning-based system for heart disease prediction and a continuous real-time cardiac health monitoring system, leveraging IoT and advanced data analytics to support early diagnosis and effective treatment.

#### **IV. OBJECTIVE OF THE PAPER**

The objective of this research paper is to comprehensively explore the application of Machine Learning (ML) in healthcare, addressing the following key research objectives:

1. Investigate the necessity of ML in healthcare, emphasizing its role in enhancing patient care quality and optimizing healthcare facilities.
2. Explore the diverse features associated with ML to bolster the framework of healthcare systems, enabling healthcare professionals to work more efficiently and make informed decisions in real-time.
3. Examine the fundamental pillars of ML specifically tailored for healthcare, focusing on their adaptability and effectiveness in analyzing large volumes of medical data to identify anomalies, patterns, and trends, thus minimizing human error.
4. Identify and discuss significant applications of ML in healthcare settings, showcasing how ML models, through accurate analysis and efficient computing capabilities, contribute to improved patient outcomes and provide valuable medical insights previously inaccessible.<sup>[4]</sup>

The proposed research aims to merge theoretical exploration with practical validation, leveraging techniques such as adjusting classifier parameters based on performance metrics and validating approaches in established medical

databases like the MIT BIH Arrhythmia Database. Furthermore, it seeks to demonstrate how ML facilitates the validation of medical decisions and reasoning through predictive algorithms, thereby enhancing the overall quality of patient care.<sup>[5]</sup>

## V. METHODS

### Experimental Study Overview

The experimental study constitutes a multifaceted approach involving meticulous dataset preparation, intricate feature extraction and selection techniques, and sophisticated machine learning-based classification methodologies. The overarching objective of this study is to discern and categorize heartbeats into either normal or abnormal classes leveraging the rich repository of electrocardiogram (ECG) data sourced from the renowned MIT-BIH Arrhythmia Database and the MIT-BIH Supraventricular Arrhythmia Database<sup>[6]</sup>.

### Dataset Preparation

The dataset preparation phase is a critical preliminary step in ensuring the efficacy and robustness of subsequent analyses. It involves the acquisition and curation of high-quality ECG recordings from the aforementioned databases. The MIT-BIH Arrhythmia Database, in particular, is widely recognized for its extensive collection of annotated ECG signals obtained from various clinical settings. Each recording encompasses valuable insights into cardiac rhythm abnormalities, facilitating comprehensive research endeavors. Similarly, the MIT-BIH Supraventricular Arrhythmia Database offers a diverse array of ECG data specifically focused on supraventricular arrhythmias, enriching the dataset with additional pathological variations.

During dataset preparation, meticulous attention is paid to data integrity, completeness, and annotation accuracy. The records undergo thorough scrutiny and validation by domain experts to ensure reliability and consistency. Any discrepancies or anomalies are meticulously addressed to maintain the integrity of the dataset<sup>[7]</sup>.

### Feature Extraction and Selection

Feature extraction and selection play a pivotal role in distilling pertinent information from raw ECG signals, thereby enabling effective discrimination between normal and abnormal heartbeats. Various signal processing techniques, including but not limited to Discrete Wavelet Transform (DWT), are employed to extract discriminative features that encapsulate key characteristics of cardiac rhythm dynamics.

These extracted features encompass a diverse array of morphological, temporal, and amplitude-based attributes inherent in the ECG waveform. Morphological features capture the nuanced structural nuances of the ECG waveform, such as QRS complex morphology and duration. Temporal features encapsulate temporal relationships between successive cardiac events, including RR intervals and their ratios. Amplitude-based features focus on quantifying the magnitude and amplitude variations of specific waveform components, such as the QRS complex.

The selection of optimal features is guided by rigorous statistical analyses, dimensionality reduction techniques, and domain expertise. Only the most discriminative and informative features are retained for subsequent classification tasks, ensuring maximal predictive performance while mitigating the curse of dimensionality<sup>[8]</sup>.

### Machine Learning-Based Classification

Machine learning serves as the cornerstone of the classification pipeline, enabling automated and data-driven decision-making processes based on extracted features. Various supervised learning algorithms, including ensemble methods like Gradient-Boosted Trees (GBT) and Random Forests (RF), are harnessed to discern subtle patterns and correlations within the feature space<sup>[9]</sup>.

These algorithms leverage the annotated ECG dataset to learn complex mappings between input features and corresponding heartbeat classes. Through iterative training and validation processes, the models iteratively refine their predictive capabilities, thereby enhancing their ability to accurately classify unseen ECG recordings.

The performance of the classification models is rigorously evaluated using standard metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC). Extensive cross-

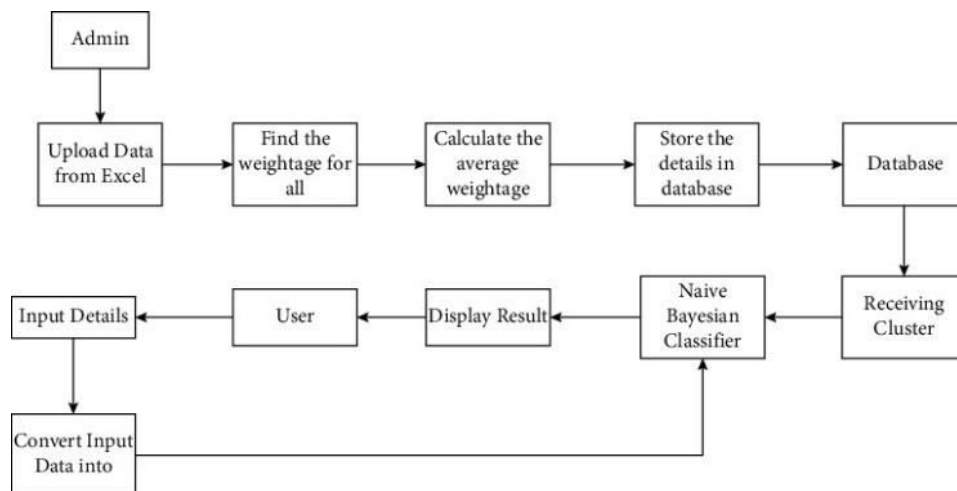
validation procedures and robustness analyses are conducted to assess the generalizability and reliability of the trained models across diverse patient populations and pathological conditions..<sup>[10][11]</sup>

### Heartbeat Dataset Preparation

A comprehensive dataset of digital ECG records was sourced from the Massachusetts Institute of Technology (MIT) arrhythmia database, specifically the MIT-BIH Arrhythmia Dataset. This dataset was accessed via the Physionet ATM Bank<sup>[8]</sup>. Each ECG record was fully extracted, resulting in a dataset with 205,146 rows from 51 patients, each row containing 16 columns of features. These records included three types of files: .atr (annotations), .hea (headers), and .dat (digital signals), digitized at 360 samples per second per channel with 11-bit resolution over a 10 mV range. Two or more cardiologists independently reviewed the dataset annotations to ensure accuracy. The processed data was saved in CSV format for further analysis<sup>[12]</sup>.

### Feature Extraction and Selection

Features were extracted using the Discrete Wavelet Transform (DWT) from the continuous ECG signal, focusing on three main types: summits features, temporal features, and morphological features. These features collectively offer insights into the ECG signals' amplitude, time intervals, and cross-correlation values<sup>[13]</sup>.



- Summits Features: These include the amplitudes of the QRS complex, representing key points in the ECG signal.
- Temporal Features: Nine features were derived, focusing on the time intervals between successive QRS peaks, such as the RR intervals, and their ratios.
- Morphological Features: These features involve the cross-correlation values between beats and the duration of QRS complexes.

In total, 16 features were identified and used for classification purposes.

### Heartbeat Classification Using Machine Learning

The classification process involved several steps:

1. Features Acquisition and Storing: Features were obtained using Matlab and stored in a CSV file. The dataset was processed using Spark-Scala, which is more efficient for handling large datasets compared to Matlab<sup>[14]</sup>.
2. Processing: The data processing steps included:
  - Filling null values with zero to prevent misclassification.
  - Labeling the heartbeat type as either normal or abnormal.
  - Addressing overfitting by balancing the dataset.
  - Separating columns by type (integer, double).
  - Using a string indexer for labeling the beat type<sup>[15]</sup>.

3. Classification Algorithms: Two machine learning models were implemented:

- Gradient-Boosted Trees (GBT): An ensemble method for regression and classification, optimized using parameters like MaxDepth and MaxIter.
- Random Forest (RF): An ensemble method utilizing decision trees, optimized with parameters like Number of Trees and MaxDepth<sup>[16][17]</sup>.

#### Gradient-Boosted Trees Model

The GBT model was built and tested with various values for MaxDepth and MaxIter to optimize performance<sup>[18]</sup>.

#### Random Forest Model

Similarly, the RF model was constructed and evaluated with different values for Number of Trees and MaxDepth to achieve optimal results<sup>[19]</sup>.

### **Study on Heart Disease Detection and Monitoring**

#### Data Collection and Algorithm Selection

Different heart disease datasets were collected to train various machine learning algorithms using the Java-based WEKA platform. The algorithm with the best performance, based on accuracy and reliability, was chosen to develop a cloud-based heart disease prediction mobile application<sup>[20]</sup>.

#### Application and Monitoring System Design

- Mobile Application: An Android application was designed where patients can input heart disease parameters manually or via sensors to predict heart disease. The application allows for report sharing with doctors and live video consultations.
- Cloud Server: The application stores doctor and patient information on a cloud server. Real-time sensor data (heartbeat, temperature, humidity) is automatically updated every 10 seconds.
- Alert System: An Arduino-based wireless monitoring system was developed to collect real-time physiological data. Critical conditions trigger alerts via GSM module and buzzers in ICUs<sup>[21][23]</sup>.

#### Continuous Cardiac Monitoring System

The system design involves real-time data collection using Arduino microcontrollers equipped with sensors. Alerts are sent to doctors via SMS in case of critical conditions, and an alarm system alerts caretakers in ICUs. Doctors can access patient data and provide consultations through the mobile application<sup>[22]</sup>.

### **Detailed Workflow and Design**

1. Collection and Selection of Datasets: Various datasets pertaining to heart diseases were gathered to train machine learning models. This included assessing different datasets for their relevance and quality.
2. Algorithm Comparison: Multiple data mining algorithms were evaluated for their performance in predicting heart disease. This step involved comparing algorithms based on metrics such as accuracy, precision, recall, and F1 score.
3. Selection of Optimal Algorithm: Based on the performance comparison, the algorithm with the highest accuracy and reliability was chosen. This algorithm was then used to develop a cloud-based intelligent heart disease prediction mobile application.
4. Application Development: The mobile application was designed to allow patients to input heart disease parameters manually or through sensors. Patients can predict their heart disease status and share results with doctors registered on the application.
5. Cloud Server Integration: All data, including patient and doctor information and real-time sensor data, is stored on a cloud server. This ensures that data is easily accessible and up-to-date.

6. Real-Time Monitoring System: An Arduino-based system was developed to monitor patient health in real-time using sensors for temperature, humidity, and heartbeat. This system updates the cloud server every 10 seconds and sends alerts via SMS to doctors if critical values are detected.

7. Alert and Notification System: In case of critical health conditions, the system sends an alert message containing current sensor data to the doctor's phone. The doctor can then log into the application, review the patient's historical data, and provide necessary medical advice through live video streaming.

8. ICU Alert System: For ICU patients, an alarm system with a buzzer alerts caretakers or nurses immediately in case of critical conditions.

By integrating these detailed methods and processes, the study aims to provide a robust and efficient system for heartbeat classification and heart disease detection, leveraging advanced machine learning techniques and real-time monitoring capabilities.

## VI. RESULT ANALYSIS

### Machine Learning Algorithm Evaluation on a Dataset

This study evaluated the performance of various machine learning algorithms on a dataset. Out of more than 10 algorithms applied, only the 5 most promising ones (with accuracy exceeding 80%) were further analyzed.

### Evaluation Method

Weka software (version 3.8.2) was used to apply 10-fold cross-validation for performance measurement. This technique helps reduce overfitting and provides a more robust estimate of the algorithms' generalizability.

### Performance Measures

The following metrics were used to assess the algorithms' performance:

$$\text{Precision} = \text{tp} / (\text{tp} + \text{fp})^{[24]}$$

$$\text{Recall} = \text{tp} / (\text{tp} + \text{fn})^{[25]}$$

$$\text{F-score} = 2 * (\text{precision} * \text{recall}) / (\text{precision} + \text{recall})^{[26]}$$

$$\text{Accuracy} = (\text{tp} + \text{tn}) / n^{[27]}$$

$$\text{Sensitivity} = \text{tp} / (\text{tp} + \text{fn})^{[28]} \text{ (same as Recall)}$$

$$\text{Specificity} = \text{tn} / (\text{tn} + \text{fp})^{[29]}$$

Where:

n = Total number of instances

tp = True positive

tn = True negative

fp = False positive

fn = False negative

These metrics are defined using true positives (tp), true negatives (tn), false positives (fp), and false negatives (fn). The formulas for each metric are provided.

## VII. CONCLUSION

In conclusion, this study has successfully validated the ability to classify heartbeats using a combination of features and machine learning classification algorithms. The developed Gradient Boosting Trees (GBT) and Random Forest (RF) models demonstrate promising results in classifying various ECG heartbeat types, potentially aiding in the quick and reliable diagnosis of cardiac arrhythmia. These models hold the potential for implementation into Computer-Aided Diagnosis (CAD) ECG systems, providing valuable support to cardiologists in interpreting ECG signals and enhancing understanding of cardiac health.



Furthermore, future endeavors could explore expanding the classification capabilities to accommodate a wider range of heartbeat types and integrating real-time data transmission from Photoplethysmography (PPG) based blood pressure sensors or electronic sphygmomanometers. The incorporation of such technologies would enrich the patient monitoring system, contributing to more comprehensive health assessment and disease prediction. Additionally, the development of a dedicated server and database tailored for this application could enhance its scalability and accessibility, potentially leading to its availability on platforms like the Android Play Store.

It's worth noting the ongoing research in PPG-based blood pressure monitoring and the broader landscape of digital health technologies, indicating a promising trajectory for the integration of advanced monitoring solutions into clinical practice. As technology continues to evolve, the potential for real-time disease prediction and remote patient monitoring becomes increasingly feasible, offering significant benefits in terms of early intervention and improved healthcare outcomes.

In summary, this work represents a significant step towards leveraging machine learning and sensor technologies for enhanced cardiac healthcare. By bridging the gap between data-driven analysis and clinical practice, these advancements hold the promise of revolutionizing the way we diagnose and manage heart-related conditions in the future.

#### REFERENCES

1. World Health Organization (WHO). "Cardiovascular Diseases (CVDs)." Available at: [www.who.int](http://www.who.int).
2. World Heart Federation. "Deaths from cardiovascular disease surged 60% globally over the last 30 years: Report." 20 May 2023. Available at: [www.world-heart-federation.org](http://www.world-heart-federation.org).
3. National Institutes of Health (NIH). "Heart Disease Facts." Available at: [www.nih.gov](http://www.nih.gov).
4. Mohd Javaid, Abid Haleem, Ravi Pratap Singh, Rajiv Suman, Shanay Rab "Significance of machine learning in healthcare: Features, pillars and applications"
5. Fajr Ibrahim Alarsan and Mamoon Younes "Analysis and classification of heart diseases using heartbeat features and machine learning algorithms"
6. Goldberger, A. L., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, 101(23), E215-E220.
7. Moody, G. B., & Mark, R. G. (2001). The impact of the MIT-BIH Arrhythmia Database. *IEEE Engineering in Medicine and Biology Magazine*, 20(3), 45-50.
8. Daubechies, I. (1992). Ten lectures on wavelets. SIAM. Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering*, (3), 230-236.
9. Zaharia, M., et al. (2010). Spark: Cluster computing with working sets. *HotCloud*, 10(10-10), 95.
10. MIT-BIH Arrhythmia Database. (n.d.). Retrieved from <https://physionet.org/content/mitdb/1.0.0/>
11. MIT-BIH Supraventricular Arrhythmia Database. (n.d.). Retrieved from <https://physionet.org/content/svdb/1.0.0/>
12. Goldberger, A. L., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, 101(23), E215-E220.
13. Moody, G. B., & Mark, R. G. (2001). The impact of the MIT-BIH Arrhythmia Database. *IEEE Engineering in Medicine and Biology Magazine*, 20(3), 45-50.
14. Daubechies, I. (1992). Ten lectures on wavelets. SIAM.
15. Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. *IEEE Transactions on Biomedical Engineering*, (3), 230-236.
16. Zaharia, M., et al. (2010). Spark: Cluster computing with working sets. *HotCloud*, 10(10-10), 95.
17. Malik, M. (1996). Heart rate variability. *Annals of Noninvasive Electrocardiology*, 1(2), 151-181.
18. Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232.
19. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
20. Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785--794). Association for Computing Machinery.
21. Han, J., & Kamber, M. (2011). *Data Mining: Concepts and Techniques*. Elsevier.

22. Patel, V., et al. (2012). A review on real-time monitoring system for automatic detection of cardiac diseases using heartbeat classification. *International Journal of Computer Applications*, 54(7), 36-42.
23. Arun, M. D., et al. (2019). Real-time Patient Monitoring System Based on Internet of Things. *International Journal of Engineering and Advanced Technology*, 8(6), 2716-2721.
24. Hazra, A., Mandal, S., Gupta, A. and Mukherjee, A. (2017) Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review. *Advances in Computational Sciences and Technology*, 10, 2137-2159.
25. Patel, J., Upadhyay, P. and Patel, D. (2016) Heart Disease Prediction Using Machine learning and Data Mining Technique. *Journals of Computer Science & Electronics*, 7, 129-137.
26. Chavan Patil, A.B. and Sonawane, P. (2017) To Predict Heart Disease Risk and Medications Using Data Mining Techniques with an IoT Based Monitoring System for Post-Operative Heart Disease Patients. *International Journal on Emerging Trends in Technology (IJETT)*, 4, 8274-8281.
27. Weng, S.F., Reys, J., Kai, J., Garibaldi, J.M. and Qureshi, N. (2017) Can Machine-Learning Improve Cardiovascular Risk Prediction Using Routine Clinical Data? *PLoS ONE*, 12, e0174944. <https://doi.org/10.1371/journal.pone.0174944>
28. Zhao, W., Wang, C. and Nakahira, Y. (2011) Medical Application on Internet of Things. *IET International Conference on Communication Technology and Application (ICCTA 2011)*, Beijing, 14-16 October 2011, 660-665.
29. Chiuchisan, I. and Geman, O. (2014) An Approach of a Decision Support and Home Monitoring System for Patients with Neurological Disorders Using Internet of Things Concepts. *WSEAS Transactions on Systems*, 13, 460-469.
30. American Heart Association. (2022). Electrocardiogram (ECG or EKG). Retrieved from <https://www.heart.org/en/health-topics/heart-attack/diagnosing-a-heart-attack/electrocardiogram-ecg-or-ekg>
31. American College of Cardiology. (2024). Smartphone-Based Recognition of Heart Failure. Retrieved from <https://www.acc.org/latest-in-cardiology/articles/2024/05/20/01/42/smartphone-based-recognition-of-heart-failure>
32. Khunte, A., et al. (2023). Accurate Diagnosis of Left Ventricular Systolic Dysfunction Using Single-Lead Electrocardiograms from a Wearable Device. *npj Digital Medicine*, 6, 124.
33. *Journal of Big Data*. (2024). Optical Electrocardiogram Based Heart Disease Prediction Using Hybrid Deep Learning. *Journal of Big Data*. Retrieved from <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-024-00698-9>



INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details