



Artificial Intelligence Based Prognosis of Cardiovascular Abnormality using Photoplethysmography (PPG) Sensors in Smart Watches.

Mir Mehdi Ali Jafri

SICSR, Symbiosis International University, Pune, India

ARTICLE HISTORY

Received on: 20/04/2025

Accepted on: 01/05/2025

Available Online: 28/07/2025

Key words:

Prognosis, Artificial Intelligence, Machine learning, Deep Learning, Photoplethysmography (PPG), Sensors, Cardiovascular heart disease, Abnormality, Smart Watches.

ABSTRACT

Purpose: This study aims to address the critical issue of delayed detection of cardiovascular heart diseases, a leading cause of sudden death and heart failure in younger populations. It seeks to develop an innovative, personalized, and ambulatory system for early detection and timely intervention to prevent life-threatening outcomes.

Design/Methodology/Approach: The research proposes a novel approach utilizing Photoplethysmography (PPG) integrated into smartwatches and wearable devices. An artificially intelligent prognosis system, combining Convolutional Neural Network – Long-Short Term Memory (CNN-LSTM) and random forest algorithms, is employed to detect and quantify cardiovascular abnormalities. Datasets such as MIMIC-IV and PPG-DaLiA are used to train and validate the framework, ensuring continuous monitoring during daily activities.

Findings: The CNN-LSTM and random forest-based system effectively detects cardiovascular abnormalities with high accuracy, quantifying the degree of irregularity. The framework demonstrates the potential for real-time, non-invasive monitoring, enabling early identification of cardiovascular issues before they escalate.

Research Limitations/Implications: The study relies on existing datasets, which may not fully represent diverse populations or real-world conditions. The accuracy of PPG-based monitoring may be affected by factors such as motion artifacts or device placement. Further research is needed to validate the system across broader demographics and refine its robustness for clinical adoption.

Practical Implications: The proposed system offers an affordable, privacy-conscious, and accessible solution for continuous cardiovascular monitoring. By integrating with widely used wearable devices, it empowers individuals to proactively manage their heart health without requiring frequent medical consultations or specialized setups.

Originality/Value: This research introduces a pioneering approach to cardiovascular monitoring by combining PPG with advanced AI techniques (CNN-LSTM and random forest) in wearable technology. Its emphasis on privacy, affordability, and real-time detection during daily activities adds significant value to preventive healthcare, addressing the critical gap in early diagnosis of cardiovascular diseases.

**Corresponding Author*

Mir Mehdi Ali Jafri, Symbiosis

International University, Pune, India

Email: mirmehdialijafri@gmail.com

INTRODUCTION

Cardiovascular heart-related diseases are a major cause of death. Prediction of cardiovascular heart disease well before is an advantage, but it is very challenging to predict and diagnose to initiate a control strategy to reduce the adverse effects. (Islam et al., 2020). It is seen that millions of lives are lost due to cardiovascular heart diseases only. (Schmidt et al., 2015) (Romiti et al., 2020) (Islam et al., 2020) (Bae et al., 2022). A person should understand and see his/her internal organs' ability to perform. A healthy life is considered a basic right of every human being (WHS- 2022). The ambulatory wearable gadgets can help to get the ideal attainable level of health.

Normally, a patient reaches out to a medical expert after a cardiac arrest, stroke, or huge pain; in many cases, it becomes difficult or impossible for medical experts to save the life of a subject (Singh et al., 2024; Macfarlane et al., 2015). After doing a literature survey, it is also acceptable that early detection of heart diseases is an advantage. (Singh et al., 2024, Macfarlane et al., 2015). Every prediction and diagnosis algorithm used in various instruments provides the analysis report after the patient is completely ill. There is no device or system to see or diagnose the disease well before it reaches the above-normal level and shows the total health scorecard or degree of abnormality. Hence, early detection of cardiovascular heart disease is an advantage.

Nowadays, lots of smart wearable devices are available that can help a person to observe the body's internal and external activities as the number of steps walked, calories burned, oxygen saturation,

tachycardia, bradycardia, and atrial fibrillation (Seshadri et al., 2020).

TREATMENT GAP

The non-communicable diseases, including cardiovascular disease (CVD) and other non-communicable diseases, are major causes of death, such as Ischemia. The author has presented nine risk factors associated with cardiovascular disease. The last major risk factor is the treatment gap due to guidelines. (Kumar & Sinha, 2020) Continuous monitoring for a longer period gives useful information about cardiac activities and presence of abnormality. The sensor can monitor the electrical activities of the subject and transmit using Bluetooth; the waveform can be visualized on a mobile screen. This system can monitor logs and various types of signals, including ECG, and send alerts if any abnormality is detected; the alert message can be sent to the patient, clinicians, or any related person using IT infrastructure. The subject can monitor their self-health status. This paper is aligned with my research studies and can be a vision for developing such an accurate and self-managed ambulatory system. (Hu et al., 2012)

Diagnosis outside the medical setup is not available: Diagnosis is usually performed in a medical setup using ECG / MRI / CT Scan. Healthcare equipment outside the medical setup is not available to the common man. Available equipment is expensive, and diagnostic equipment is critical to operate and interpret the output. (Malakuti, 2023).

Cardiac activity indicators: Health is a common issue for every human being. Disease and problems do not occur suddenly, but they are continuous processes. If a health problem starts, it also

gives basic and common indications (Wu et al., 2003). Generally, we avoid finding the reason behind feeling uncomfortable, mild pain, giddiness, vertigo, or tiredness.

The World Health Organization uses the term Health Indicators. Cardiovascular disease is at the top among four NCDs (non-communicable diseases) and is a major risk factor. Cardiovascular diseases killed 17.9 Million people in the year 2019. (WHS, 2023).

Heartbeat patterns and sound change as time progresses, it is a natural phenomenon that new objects and old objects make different sounds after some time when used for a longer time similarly heart muscles are responsible for the changes in sound and electrical activity associated with the heart. (Macfarlane et al., 2015)

Arrhythmia: In medical terms, arrhythmia describes an abnormal or irregular heartbeat. The heart of a normal person beats in a rhythmic pattern regularly at a consistent rate and forms a regular interval between heartbeats. Electrical signals generated by the heart may malfunction and will be called arrhythmia.

Types of arrhythmias: Atrial Flutter, Bradycardia, Tachycardia, Supraventricular Tachycardia (SVT), Ventricular Tachycardia (V. Tach), Atrial Tachycardia (A. Tach.), Ventricular Fibrillation (V. Fib.), Premature Contractions, Long QT Syndrome (LQTS), Wolff-Parkinson-White (WPW) Syndrome, Sick Sinus Syndrome, and Conduction Block.

Electrocardiogram: It is known as ECG or EKG, a very common and reliable non-invasive device to see heart's electrical activities. The electrocardiogram works on

a simple natural phenomenon that heart muscle cells generate electrical current while beating. The electrocardiogram first detects the electric current, amplifies, and captures.

Output of ECG: It represents the heart rate (HR), Rhythm, abnormalities, and heart damage.

Treadmill Test (TMT): It is also known as a treadmill stress test and is a medical procedure used to assess the cardiovascular health and fitness of an individual. It shows an analysis of how the heart responds to physical stress. The primary purposes of a treadmill test are to analyze cardiac functions and assess fitness level, TMT tests are also used to determine a person's overall cardiovascular fitness level. The goal is to increase heart rate and monitor response to the exercise. In the TMT assessment, the test keeps going until the subject reaches a target heart rate, exhibits symptoms like shortness of breath or chest pain, or the doctor deems it necessary to halt it for patient safety. Results can reveal vital details of the heart. The existence of irregular cardiac rhythms, decreased heart blood flow, and other symptoms brought on by exercise.

Issues and limitations with ECG / EKG: ECG signal processing is the most critical task when diagnosing cardiovascular disorder, so the classification and the approaches based on machine learning are the primary objectives of the research. The PhysioNET initiated the automatic ECG classification using data from many sources. (Malakuti, 2023)

The electrocardiogram heart monitoring systems with microcontroller units (MCUs) can reduce the cost by 20 times. However, an efficient algorithm is essential to

develop the industry-grade system. Training machine learning models is possible in MCUs. Expensive biomedical equipment increases the treatment cost, and affordable ECG would make treatment accessible to more people and provide better treatment. (Zishan, 2024),

Any disorder that compromises a person's health is commonly called heart disease. Heart conditions include arrhythmias and ischemic heart disease, among others. The ECG signal can be used to categorize these illnesses. The electrocardiogram (ECG) is the preferred method for assessing a person's health status and diagnosing and treating heart conditions. Different wave patterns determine the type of heart disease that a person has. While some heart conditions are fatal, others are not treated. Therefore, it is necessary to classify the various heart diseases. In the proposed work, the CNN algorithm is supposed to be implemented to classify different heart diseases. Normal and abnormal signals can be distinguished with greater ease and precision. (Gawande & Barhatte, 2017)

REVIEW OF LITERATURE

Literature Review includes computer science, clinical, technical, and medical engineering research papers.

Artificial intelligence techniques are applied to wearable computing and IoT devices, which continuously monitor a subject, and the sensors are helpful to self-health monitoring outside medical settings and reduce dependency. The performance of the data acquired by smart and wearable sensors is compared with classical machine learning algorithm-based device datasets. (Huang et al., 2022)

A study performed to determine the accuracy of non-invasive monitoring from remote locations and predicting heart

failure. Invasive cardiac sensors have shown great favourable promises. Machine learning technology was used to detect heart failure. (Stehlik et al., 2020) but usually ECG recording is done in a medical setup with 12 lead ECG machines, but the author has demonstrated that handheld 6 leads ECG recording is compatible with mobile phones and transferred the cardiac recording at a remote place. They have recruited 21 patients from the cardiac ward aiming to capture ECG. (Mercer et al., 2020)

There are highly effective treatments available, but the death burden is huge. Technology inventions are promising for further quality and experience related to cardiovascular heart disease treatment. This paper provides an overview of available technologies in current practice. It provides the challenges in implementation in the European region. This paper also recaps how these challenges could be tackled. The author has discussed physician's challenges and barriers to digital health implementation legal and ethical issues, interoperability, economic issues, the role of patient education programs, the role of data standardization, the assurance, and the role of professional organizations. (Frederix et al., 2019)

An ECG analysis done at three different centers (IIT Roorkee, NIT Jalandhar, and SGGS Institute of Eng & Technology, Vishnupuri, Nanded). The cardiac electrical activity of all participants undertaken using 12 lead ECG, a total of 963 healthy individuals belonging to the southern part of India were considered whose age was between 18 to 83 years, there were 30.4% females. Normal limits were longer in males than females. These normal cardiac electrical values can be used to compare the present subject and indicate the degree of

abnormality and timely action if the indicator shows severity. (Macfarlane et al., 2015)

A Smartphone camera is used for recording of a fingertip for continuous monitoring. Smartphones can monitor and differentiate normal rhythms and premature ventricular contractions. It can improve patients' health and reduce the cost. Mobile monitoring devices are inexpensive and simple to use. They have proposed the algorithm for iPhone 4S/5S which can distinguish AF, NSR, PACs, and PVCs. (Chong et al., 2015)

The normal values for the adult, for all age groups of males and females. 13,354 individuals taken from the Netherlands, the age is ranging from 16 to 90 years in which 55% were men, 45 were women and the individuals were cardiologically healthy as per commonly accepted symptoms. The 12 - lead ECGs were used. The validated computer program was used to process all ECGs. This study is different from other studies because they have included a large population of males and females. This research paper is useful to compare with country-specific ECG patterns. I have Indian and Chinese populations with normal ECG values; it will be helpful to understand the ECG pattern and differences between geographical populations. (Rijnbeek et al., 2014)

The ECG normal limits are the basis for diagnosis. The electrocardiogram considers the subject's age and gender to determine heart functionality. The study shows normal limits in Chinese subjects, they have considered a huge sample size of healthy subjects. The methods were standard. The results were significant age trends in the Chinese population. This normal value will be used to compare the current subjects' acquired data and predict the degree of

abnormality and other indications. (Wu et al., 2003)

RESEARCH GAP

It is clear from the literature review that heart diseases usually do not occur suddenly but gradually increase and reach a noticeable peak level. Once cardiovascular heart disease reaches the mature level, then it becomes difficult to revert.

Need of Continuous Monitoring:

Normally, a patient reaches out to a medical expert after a cardiac arrest, stroke, or huge pain; in many cases, it becomes difficult or impossible for medical experts to save the life of a subject (Singh et al., 2024; Macfarlane et al., 2015).

There is limited used of AIML technology to diagnose disease in initial stage:

Experts mention that diagnosing beforehand is advantageous, and detecting heart disease in earlier stages is considered a significant task in medical science (Singh, et al. 2024).

Continuous monitoring increases data and challenges:

Active prevention of heart-related diseases is preferable to passive treatment. Long-term and continuous monitoring can improve early diagnosis using battery-powered portable wearable technologies that can detect abnormal heartbeats, but battery-powered ECG acquisition degrades the signals due to compounded motion artifacts and faces several interpretation challenges by human or machine learning techniques. The denoising is applied in preprocessing using Kalman filtering, furrier decomposition, wavelet domain denoising, and EMD. (Li, et al., 2023)

The wearable sensor generates continuous data; hence, data size constantly increases due to continuous monitoring of the human

body internally and externally (Ali et al., 2022)

Data is Unstructured: The Electronic Medical Records (EMRs) are unstructured. (Ali et al., 2022) This means there are images, X-rays, ECG graphical reports, genetic information, and blood test reports in text data, so it becomes challenging to handle heterogeneous data types. The attention mechanism lets one focus dynamically on a specific relevant input data segment. (Prakash et al., 2024)

Data gets corrupted.: The EMRs are unstructured, and wearable sensor data is corrupted due to signal artifacts and continuous monitoring. Missing values and noise decrease the performance and give inaccurate results. (Ali, 2022)

Limited availability of technologies.: Machine learning algorithms ANN, CNN, SVM, Logistic regression, and Naïve Bayes play significant roles in the accurate and timely detection of heart diseases. These machine learning technologies can differentiate a normal person and a post-stroke person. (Singh et al., 2024)

The decision tree showed 91.56% accuracy, and the random forest showed 97.51% accuracy when using PPG while walking elderly people. (YU et al., 2022)

The medical field follows a scientific approach, but the prediction results are unsatisfactory. The predictions are important in the medical field. An ANN can help physicians to predict disease. The ANN algorithm's accuracy is 84.73%, so ANN and other neural network technologies can be considered reliable tools. The implementation of ANN is performed using MATLAB and its tools; since medical information is in huge amounts, physicians need an automated, accurate, and reliable system for

predictions and correct decision-making, so, in this regard, ANNs are reliable and powerful tools. A computer application is used to gather information and predict the chances of heart disease in the next 10 years. This application takes 15 input parameters and uses a trained set. This paper is aligned with my research studies and strengthens the need for early detection of disease before reaches the above degree of abnormality.

Diagnosis outside the medical setup is not available: The research work was done to evaluate AW4 to measure and compare. The wearable devices which are outside the medical setting are questioned to prove accuracy; the researcher has successfully compared the accuracy using the AW4 wearable wristwatch. Fifty post-cardiac surgery patients who were experiencing common cardiac arrhythmias like atrial fibrillation (AFib) were included in the study. Included patients were above the age of 18 years. Cardiac pacemaker patients were excluded. Bypass surgery patients and tattoo skin patients were not included. The collected data is compared with the telemetry data where study results show a correlation coefficient of 0.7 between readings and it is found in the study that AW4 watch was measuring HR accurately in patients who were having AFib and those who were not having AFib complained the correlation coefficient. This study helps a common person to measure HR continuously outside the medical settings independently. It can be used as one of the indicators to take timely action. (Seshadri et al., 2020),

Early Detection but No degree of abnormality: After doing a literature survey, it is also acceptable that early detection of heart diseases is an advantage.

(Singh et al., 2024, Macfarlane et al., 2015). Every prediction and diagnosis algorithm used in various instruments provides the analysis report after the patient is completely ill. There is no device or system to see or diagnose the disease well before it reaches the above-normal level and shows the total health scorecard or degree of abnormality.

Normal values of the heart's electrical activities are presented by (Macfarlane et al., 2015) and it can be used to benchmark and compare with the available subject to get the degree of abnormality and prepare the scorecard.

In a survey study to detect irregular pulses and or AFib using smartwatches and their Apps, a smartphone (Apple iPhone) was used by subjects who did not have atrial fibrillation, as reported by the subjects. The surveys were conducted for several days to capture irregular behaviour of pulse, notify, and show an ECG. They recruited 419,297 participants over a period of 8 months, 34% of whom had AFib on subsequent ECG patches. This study was carried out with subjects who had never visited the site. (Pérez et al., 2019)

Risk calculator unavailability: A study was carried out to evaluate the Framingham Risk Score (FRS) for the assessment of CAD. This paper becomes important because it gives the risk score of CVD in the coming 10 years which indicates that patients to take timely preventive measures. It can help the study to predict heart disease before it reaches its peak and over abnormality level. (Jahangiry et al., 2017)

Research carried out to test the score of the Framingham and the online Adult Treatment Panel- III (ATP-III) risk estimation for 10-year among younger

adults of age between 18-39 years. A total 10551 male participants were included from 1967 to 1973. The mortality rate for 30 years of follow-up, the FRS remained at 10% in the age group 18 to 29 years and increased to 12% in 30 to 39 years group. This paper helps prove that the Framingham Risks Score is an important indicator for calculating heart-related disease. (Berry et al., 2007)

PPG as an ambulatory substitute: A PPG device accurately measured several times the heartbeat (HR - Heart Rate) and several times breathed per minute (i.e., RR - Respiratory Rate). The HR and RR are usually measured with specialized equipment; measuring HR and RR is difficult outside medical settings, but in this research paper, photoplethysmography and a smartphone camera are used; the study shows that smartphones and gadgets can be used to measure HR and RR. In my opinion, if smartphones, smart gadgets, and IoT can measure the HR and RR, then this technique can be very useful to monitor HR, RR, and cardiac activity continuously and further store collected data, find patterns, and timely indicate the degree of abnormality. (Bae et al., 2022)

A model based on optical and physiological considerations where only one light source was assumed. Contactless monitoring with Remote Photoplethysmography (rPPG) makes it possible to capture human cardiac activities by detecting color variations on human skin using the multi-wavelength camera. The skin reflection model and have analyzed the strengths and weaknesses of existing rPPG methods. First, a large video dataset is introduced, and next, an evaluation matrix is prepared consisting of 8 rPPG methods for comparison. In my assumption, contactless and continuous

monitoring techniques are useful to my study and support the self-health assessment outside the medical setup and hospital and timely contact experts. (Wang et al., 2017)

The predictions of heart-related disease using ANN and various variables show heart disease and predict a subject's heart disease, data is collected using three smart tools a smart mirror, a mouse, and a phone. The data is collected for a period of one year on a server connected to the internet to get enough data to make predictions in one year which will increase awareness of self-heart and reduce death due to heart disease, but reality shows that systems are not completely integrated with such tools. Analysis is happening but predictions one year ahead are not yet done. The discussed smart tools can be used in daily life to monitor and make judgments about personal predictions. This paper is very important since my study also focuses on Ambulatory monitoring of cardiac activities outside the medical setup. (Wijaya et al., 2013)

Need for Algorithms to detect heart diseases using PPG: There are various competent techniques and algorithms used to predict cardiovascular heart disease using ECG but there are no competent techniques and algorithms to predict cardiovascular heart diseases using PPG before stroke and heart attack. According to (Ghosal et al., 2017), the ECG captures the activities of the muscles of the heart in the form of waves used to understand the abnormality and reach a conclusion but in ambulatory observation and diagnosis the signals get corrupted, and analysis of signals becomes difficult. In this research study author applies the SOM- Self Organizing Map (an unsupervised learning

and classifying technique) to remove high-frequency noise, baseline wander noise removal, and feature Extraction and finally compares with the ANN (Artificial Neural Network) to preprocess the noisy signals. The SOM can differentiate data. This study shows that the SOM is a more suitable technique in medical signal processing as compared to ANN and other techniques used.

The Self-Organizing Map (SOM) of ANN is an unsupervised machine learning technique to classify input data in groups according to its normal and abnormal features. (Ghosal et al., 2017). The SOM classifies the beats according to their shape and will keep similar beats in groups. So, by this method, it can be easier to classify corrupted beats from the ECG signal. (Ghosal et al., 2017)

Need for AIML based wearable computing system to monitor continuously: The wearable sensors generate continuous data, and the data size increases (Ali et al., 2022). The wearable sensor data gets corrupted due to signal artifacts, missing values, and noise. Noise decreases the performance and shows inaccuracy, so, an intelligent framework is needed that can automatically fuse the information from the EMRs and the sensor's data. (Ali et al., 2022)

A physicians can use big and continuously developed AI techniques to apply them according to needs, Establishing new connections between various stakeholders like patients and physicians. Physicians should use AI in cardiology heart disease diagnosis, and treatment and should adopt it further also give assurance to physicians. (Romiti et al., 2020)

According to the author, detection of arrhythmias, segment changes, prediction, important cardiac events like stroke, monitoring ECG from wearable devices continuously, extracting hidden features, therapy guidance, time required in the treatment, and lastly, AI is effective in diagnosis, treatment, and managing heart-related activities. This paper is very useful since it gives the surety that AI is used and adopted by clinicians and medical practitioners for various diagnoses and other activities. This paper becomes important because it gives the future directions too. (Martínez-Sellés & Marina-Breyse, 2023)

The research is emphasizing that the healthcare sector produces huge amounts of data on heart disease that are not used effectively to locate secret patterns for decision-making to detect heart disease before it reaches the above-normal level. Heart disease is considered a major cause of death, and it is difficult to predict. The research work reduces the dimensionality of the dataset. (Islam et al., 2020)

RESEARCH QUESTIONS

The research questions in the current proposed research work are based on the research gaps as mentioned below. Heartbeats and electrical activity pattern as age increases.

- i. How can stroke or heart attack be identified before it occurs using AIML?
- ii. How can the degree of abnormality of the heart's activities be quantified?
- iii. How to develop an AIML based scorecard and compare it over a period?
- iv. What are the limitations of available diagnostic equipment?

- v. How can the heart function be compared with the present heart after a long time using AIML techniques?
- vi. Can we develop an AI-based prototype that indicates the heart disease or degree of abnormality well before it reaches its peak level using a health scorecard?
- vii. How can smart gadgets and IoT be used to predict a heart attack, arrhythmia, Afib, Vfib, Ischaemia, and Myocardial Infarction?

Inspiration: Smart watches and fitness bands still have the scope to become more intelligent. Smart gadgets can provide step count, HB, BP, Stress, AFib, and such information. CGM - Continuous Glucose Monitoring System. Sugar level testing electronic kit, which is a small device, takes input process and gives output used by individuals at home. Smart gadgets and IoT technology are reachable.

Palliative (comfort care) is increasing, but data is lacking. According to the authors, a structured literature review revealed no quality indicators of palliative care in heart disease. (Mizuno et al., 2017)

The above devices inspire the development of IoT-based systems that are used outside the medical setup and more importantly, operated independently to monitor continuously and check the degree of abnormality.

STATEMENT OF THE PROBLEM

The technology has changed dramatically from 300 lead recording to single lead recording, and the use of single lead, historically automated ECG interpretation began around 1960 and still improving

standards and protocols from well-defined database modelling which uses big data and Artificial Intelligence. Some equipment has built-in computer algorithms. BSM - Body Surface Mapping can be used to visualize other abnormalities. The author discusses the applications of ECGI and shows that it is still alive in ECG research. In this paper, it is sure that one lead ECG monitoring and wearable gadgets are more in demand due to reliability and accuracy. (Macfarlane, 2022)

Nowadays, lots of smart wearable devices are available that can help a person to observe the body's internal and external activities like the number of steps walked, calories burned, oxygen saturation, tachycardia, bradycardia, and atrial fibrillation (Seshadri et al., 2020)

Smartwatches and fitness bands are easily available gadgets, but there are no ambulatory continuous heart monitoring systems and gadgets that cannot show the continuous activities of the heart. ELR and Holter data analysis is not available to common populations. The computer is a highly advanced machine that can process several input types and produce the desired output. Large trials and observations can be carried out easily using the internet and other devices. (Schmidt et al., 2015).

OBJECTIVES OF THE STUDY

The Objectives of our research studies are given below.

1. To detect the outliers in time series data generated during the continuous monitoring of cardiovascular activities using PPG.

2. To quantify the degree of abnormality in cardiovascular activities.
3. To develop an indicator model to indicate severity and abnormality in the cardiovascular activities.

SCOPE AND LIMITATIONS

The scope of my research will be based on the research gaps identified and considering current limitations of the available system, gadgets, techniques, and technologies.

In ambulatory monitoring, ECG beats get corrupted, so the biomedical signal analysis becomes difficult and hence efficient processing is needed. (Ghosal et al., 2017)

Get a Secondary Dataset of PPG / ECG and use available algorithms of AIML from open-source software and libraries. Analysis and comparison of various algorithms. Develop a prototype of a health scorecard to check the degree of abnormality.

Small kids are not included in the study, so this study does not apply to minors and heart-attacked patients. This study will be performed in India and will not include people outside India.

Data source: This current research study will use secondary data from multiple sources like PhysioNET Pulse Transit Time PPG dataset (www.physionet.org), MIMIC-III & IV (26 Tables, 40000 patients data), PPD-DaLiA, PuPG, Figshare PPG-BP, Cleveland Dataset, and MIT-BIH. Since the feature fusion technique is used in the current research study, ECG, PPG, and pathology reports will be used, and a set of selected data will be considered for research work.

Variable Vital Signs:

ID**Age****Gender****BP-** Blood Pressure**BP-** Base Point**SP-** Systolic Peak**DC-** Dicrotic Notch**DP-** Diastolic Peak**HR-** Heart Rate**HRV-** Heart Rate Variability**RR-**Respiratory Rate**Time Series Data-** (Hourly, Daily, Minutes by Minutes)

METHODOLOGY, TOOLS, AND TECHNIQUES

Data Pre-processing: The data from ECG in tabular format is used, PPG is in wave format, which will be converted in tabular format using the Python programming library Wave Format Database (WFDB).

Feature fusion: It is the process of combining diverse data types and electronic medical records and considering the combination of a few selected types of features for research.

Weighted Sum in feature fusion: if x_1 and x_2 represent two features and w_1 and w_2 represent two weights then the fused feature xf will be represented as follows

$$xf = f_1w_1 + f_2w_2 +$$

Classification: Classification algorithms will predict the presence or absence of heart disease. Popular classification techniques are Navies Bayes, Decision Tree, and Support Vector Machine etc.

Methodology and algorithms recommended: To predict heart disease at an early stage; it is believed that disease and problems do not occur suddenly; it is a continuous process that starts early and

starts giving basic and common indications. After ample time, suddenly reaches a peak level, and only most people contact experts. This paper gives the conceptual hybrid model based on data science and its various algorithms to predict or indicate the severity in advance and proposes to suggest globally accepted precautionary steps. Naïve Bayes, ANN, SVM. The result shows increased accuracy, specificity, sensitivity, and increased attributes like Body Mass Index. (Junaid & Kumar, 2020)

There are two methods to detect heart rate from fingertip images. According to the authors, smartphones and their cameras are important devices to monitor heart rhythm. A mobile can monitor and distinguish atrial fibrillation. (Zaman et al., 2017)

(Banu & Swamy, 2016), reviewed 21 different research papers published between 2004 and 2016 based on data mining, the accuracy of methods, and machine learning technologies using datasets of patients showing various factors. A physician can predict and help the patient by using intelligent model in medical system.

Technology used: In this current research study, the following technologies will be the major focus including other minor technologies like AIML, SOM, LLN, PPG, ECG etc.

Artificial Intelligence & Machine Learning: Machine Learning (ML) is considered the essential part of Artificial Intelligence (AI), which emphasizes statistical models and methods of classification to find relations between objects that are used by computer systems and other devices to improve their performance to do any task based on the available data, without explicit instructions.

Self-Organizing Map (SOM): It is also known as a Kohonen map or Kohonen network and is an unsupervised learning technique in artificial neural networks. It was developed by Finnish scientist Teuvo Kohonen in the 1980s. The SOM may be used for dimensionality reduction, in other words, SOM will produce low dimensional data from the higher dimensional data. SOM can be used in visualization and clustering. It is particularly effective for representing and visualizing high-dimension data.

LNN- Liquid Neural Network (LNN) is one of the ML algorithms which deals time series dataset.

Photoplethysmography (PPG): Photoplethysmography stands for PPG, it is a non-invasive measurement method. An LED light source is used to throw light and capture the reflections to measure. PPG is frequently used to track several physiological variables in blood volume brought on by each heartbeat.

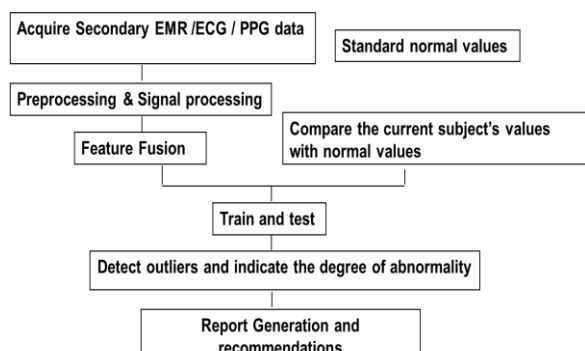
Why these technologies? (AIML, SOM, LNN, and PPG): Photoplethysmography (PPG) can be an alternative to ECG and other diagnostic technologies. The major advantage of PPG is ambulatory character; the ambulatory test, diagnosis, or observation can be done while walking and/or working, whereas other technologies require medical setup and dependency on experts. AI, ML, SOM, LLN, and PPG technologies are essential components to monitor continuously, automate the process, achieve the highest accuracy, and indicate the degree of abnormality in the starting phase only so that timely decisions can be taken to control the situation.

IoT Technology to monitor remotely: (Li et al., 2017), written a research paper, according to the author, traditional systems are passive and patients die before getting treatment because they fail to contact healthcare experts or are unaware of the effects. The author highlighted the value of IoT, the IoT device can invoke the service. The proposed pervasive (persistent) system has two layers to acquire and transmit the data. ECG, Spo2, HR, Pulse rate, blood fat, glucose, and environment indicator (location) are monitored continuously. Presented four modes 1) Patient's risk, 2) Need of medical analysis, 3) Demand for communication, 4) Computing resources. The proposed prototype sends warnings to healthcare services and concerns about potential attacks in advance. There are four important sections, System Architecture, Data Acquisition part, Operation mode for monitoring system, and Prototype Implementation.

WORK PLAN SUMMARY

1. Study Signal Processing equipment, tools, and techniques to acquire EMR/ ECG/ PPG secondary dataset.
2. Data pre-processing (EMR/ECG/PPG data).
3. Feature fusion.
4. Compare with standard values.
5. Train and test to predict the degree of abnormality.
6. Indicate the degree of abnormality.
7. Report generation

Work plan logical view (Input, Process, and Output)



Objective 1- To detect the outliers in time series data generated during the continuous monitoring of cardiovascular activities using PPG.

Materials & Methods

Threshold-Based Detection:

1. Visual Methods

Box Plot, Scatter Plot, Histogram

2. Statistical Methods

Z-Score, Modified Z-Score, Interquartile Range (IQR):

3. Machine Learning Methods

Isolation Forest, SVM

4. Domain-Specific Methods.



Figure 2: Objective 1 work plan

Objective 2- To quantify the degree of abnormality in cardiovascular activities.

Assigning a score that reflects how unusual the data point is to the rest of the values.

Statistical model: Z-Score, Percentile rank

Distance-Based Methods: Mahalanobis Distance, Euclidean Distance

Density-Based Methods: Local Outlier Factor (LOF)

Model-Based Methods: Cook's Distance.

Threshold-Based Scoring.

Combining Multiple Measures: Weighted Averaging & Ensemble Methods.

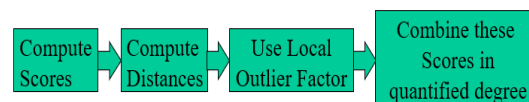


Figure 3: Objective 2 work plan

Objective 3- To develop an indicator model to indicate severity and abnormality in the cardiovascular activities.

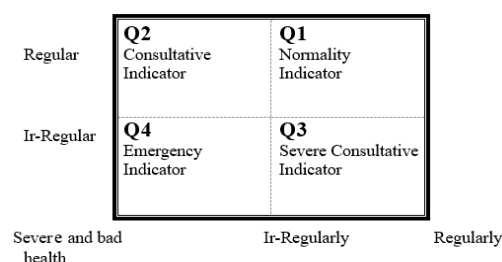


Figure 4: Abnormality Quadrants

There can be following indicators

1. Severity indicator
2. Abnormality indicator
3. Degree of normality

Q1 Normality Indicator, awareness about self-health.

Q2 & Q3 Abnormality above normal values, advises to consult cardiac expert.

Q4 Emergency Indicator advises to consult immediately



Figure 5: Objective 3 work plan

Input: EMR, PPG(bp, sp, dn, dp), ECG(PQRST), BP, intervals and segments.

Process: Process the captured data, and general identification procedure for decision-making. Classify the subject's cardiac activities. The standard normal values of EMRs will be used as an indicator to compare with the current subject's

cardiac activities and degree of abnormality, and other indicators will show normality, severity, and emergency.

Output: Indicate the degree of abnormality before it reaches the peak and irreversible level. Users can apply preventive strategies according to the score of abnormality before a heart attack or stroke. Consult the medical expert and go for the next level of diagnosis.

REFERENCES

1. Ali, F., El-Sappagh, S., Islam, S. M. R., Kwak, D., Ali, A., Imran, M., & Kwak, K. (2020). A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Information Fusion*, 63, 208–222. <https://doi.org/10.1016/j.inffus.2020.06.008>
2. Bae, S., Borac, S., Emre, Y., Wang, J., Wu, J., Kashyap, M., ... Po, M. J. (2022). Prospective validation of smartphone-based heart rate and respiratory rate measurement algorithms. *Communications Medicine*, 2(1). <https://doi.org/10.1038/s43856-022-00102-x>
3. Banu, N. K. S., & Swamy, S. (2016). Prediction of heart disease at early stage using data mining and big data analytics: A survey. In *IEEE International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT)*. <https://doi.org/10.1109/iceeccot.2016.7955226>
4. Berry, J. D., Lloyd-Jones, D. M., Garside, D. B., & Greenland, P. (2007). Framingham risk score and prediction of coronary heart disease death in young men. *American Heart Journal*, 154(1), 80–86. <https://doi.org/10.1016/j.ahj.2007.03.042>
5. Chong, J. W., Esa, N., McManus, D. D., & Chon, K. H. (2015). Arrhythmia discrimination using a smart phone. *IEEE Journal of Biomedical and Health Informatics*, <https://doi.org/10.1109/jbhi.2015.2418195>
6. Frederix, I., Caiani, E. G., Dendale, P., Anker, S. D., Bax, J. J., Böhm, A., ... Van Der Velde, E. (2019). ESC e-Cardiology Working Group position paper: Overcoming challenges in digital health implementation in cardiovascular medicine. *European Journal of Preventive Cardiology*, 26(11), 1166–1177. <https://doi.org/10.1177/2047487319832394>
7. Gawande, N., & Barhatte, A. (2017). Heart diseases classification using convolutional neural network. In *Proceedings of the 2nd International Conference on Communication and Electronics Systems (ICCES)*. <https://doi.org/10.1109/cesys.2017.8321264>
8. Ghafoori, M., Clevenger, C. M., Abdallah, M., & Rens, K. (2023). Heart rate modeling and prediction of construction workers based on physical activity using deep learning. *Automation in Construction*, 155, 105077. <https://doi.org/10.1016/j.autcon.2023.105077>
9. Ghosal, P., Sarkar, D., Kundu, S., Roy, S., Sinha, A., & Ganguli, S. (2017). ECG beat quality assessment using self-organizing map. *IEEE Explore*. <https://doi.org/10.1109/optronix.2017.8349994>
10. Hu, S., Wei, H., Chen, Y., & Tan, J. (2012). A real-time cardiac arrhythmia classification system with wearable sensor networks. *Sensors*, 12(9), 12844–12869. <https://doi.org/10.3390/s120912844>
11. Huang, J., Wang, J., Ramsey, E., Leavey, G., Chico, T. J. A., & Condell, J. (2022). Applying artificial intelligence to wearable sensor data to diagnose and predict cardiovascular disease: A review. *Sensors*, 22(20), 8002. <https://doi.org/10.3390/s22208002>
12. Islam, M. T., Rafa, S. R., & Kibria, M. G. (2020). Early prediction of heart disease using PCA and hybrid genetic algorithm with k-means. In *23rd International Conference on Computer and Information Technology (ICCIT)*. <https://doi.org/10.1109/iccit51783.2020.9392655>

13. Jahangiry, L., Farhangi, M. A., & Rezaei, F. (2017). Framingham risk score for estimation of 10-years of cardiovascular diseases risk in patients with metabolic syndrome. *Journal of Health, Population and Nutrition*, 36(1).<https://doi.org/10.1186/s41043-017-0114-0>
14. Junaid, M. J. A., & Kumar, R. (2020). Data science and its application in heart disease prediction. In *International Conference on Intelligent Engineering and Management (ICIEM)*.<https://doi.org/10.1109/iciem48762.2020.9160056>
15. Kumar, A., & Sinha, N. (2020). Cardiovascular disease in India: A 360 degree overview. *Medical Journal Armed Forces India*, 76(1), 1–3. <https://doi.org/10.1016/j.mjafi.2019.12.005>
16. Li, C., Hu, X., & Zhang, L. (2017). The IoT-based heart disease monitoring system for pervasive healthcare service. *Procedia Computer Science*, 112, 2328–2334. <https://doi.org/10.1016/j.procs.2017.08.265>
17. Li, Y., Luo, J., Dai, Q., Eshraghian, J. K., Ling, B. W., Zheng, C., & Wang, X. (2023). A deep learning approach to cardiovascular disease classification using empirical mode decomposition for ECG feature extraction. *Biomedical Signal Processing and Control*, 79, 104188. <https://doi.org/10.1016/j.bspc.2022.104188>
18. Lou, Y., Lin, C., Fang, W., Lee, C., & Lin, C. (2023). Extensive deep learning model to enhance electrocardiogram application via latent cardiovascular feature extraction from identity identification. *Computer Methods and Programs in Biomedicine*, 231, 107359.
19. Macfarlane, P. W. (2022). The application of computer techniques to ECG interpretation. *Hearts*, 3(1), 1–5. <https://doi.org/10.3390/hearts3010001>
- Macfarlane, P. W., Lloyd, S. M., Singh, D., Hamde, S. T., Clark, E., Devine, B., Francq, B., & Kumar, V. (2015). Normal limits of the electrocardiogram in Indians. *Journal of Electrocardiology*, 48(4), 652–668.<https://doi.org/10.1016/j.jelectrocard.2015.04.013>
20. Malakouti, S. M. (2023). Heart disease classification based on ECG using machine learning models. *Biomedical Signal Processing and Control*, 84, 104796. <https://doi.org/10.1016/j.bspc.2023.104796>
21. Martínez-Sellés, M., & Marina-Breyse, M. (2023). Current and future use of artificial intelligence in electrocardiography. *Journal of Cardiovascular Development and Disease*, 10(4), 175. <https://doi.org/10.3390/jcdd10040175>
- Mercer, B., Leese, L., Ahmed, N., Holden, A. V., & Tayebjee, M. H. (2020). A simple adaptation of a handheld ECG recorder to obtain chest lead equivalents. *Journal of Electrocardiology*, 63, 54–56. <https://doi.org/10.1016/j.jelectrocard.2020.10.005>
22. Mischie, N., & Albu, A. (2020). Artificial neural networks for diagnosis of coronary heart disease. 2020 International Conference on e-Health and Bioengineering (EHB). <https://doi.org/10.1109/ehb50910.2020.9280271>
23. Mizuno, A., Miyashita, M., Hayashi, A., Kawai, F., Niwa, K., Utsunomiya, A., Kohsaka, S., Kohno, T., Yamamoto, T., Takayama, M., & Anzai, T. (2017). Potential palliative care quality indicators in heart disease patients: A review of the literature. *Journal of Cardiology*, 70(4), 335–341. <https://doi.org/10.1016/j.jjcc.2017.02.010>
24. Pérez, M., Mahaffey, K. W., Hedlin, H., Rumsfeld, J. S., García, A., Ferris, T., Balasubramanian, V., Russo, A. M., Rajmane, A., Cheung, L., Hung, G., Lee, J., Kowey, P. R., Talati, N., Nag, D., Gummidipundi, S., Beatty, A. L., Hills, M. T., Desai, S., ... Investigators, A. H. S. (2019). Large-scale assessment of a smartwatch to identify atrial fibrillation. *The New England Journal of Medicine*, 381(20), 1909–1917.
25. Prakash, J. V., Vijay, S. A. A., Kumar, P. G., & Karthikeyan, N. K. (2024). A novel attention-based cross-modal transfer

- learning framework for predicting cardiovascular disease. *Computers in Biology and Medicine*, 170, 107977. <https://doi.org/10.1016/j.compbimed.2024.107977>
26. Rijnbeek, P. R., Van Herpen, G., Bots, M. L., Man, S., Verweij, N., Hofman, A., Hillege, H. L., Numans, M. E., Swenne, C. A., Witteman, J. C. M., & Kors, J. A. (2014). Normal values of the electrocardiogram for ages 16–90 years. *Journal of Electrocardiology*, 47(6), 914–921. <https://doi.org/10.1016/j.jelectrocard.2014.07.022>
 27. Romiti, S., Vinciguerra, M., Saade, W., Cortajarena, I. A., & Greco, E. (2020). Artificial intelligence (AI) and cardiovascular diseases: An unexpected alliance. *Cardiology Research and Practice*, 2020, 1–8. <https://doi.org/10.1155/2020/4972346>
 28. Schmidt, S. E., Holst-Hansen, C., Hansen, J. H. L., Toft, E., & Struijk, J. (2015). Acoustic features for the identification of coronary artery disease. *IEEE Transactions on Biomedical Engineering*, 62(11), 2611–2619. <https://doi.org/10.1109/tbme.2015.2432129>
 29. Seshadri, D. R., Bittel, B., Browisky, D., Houghtaling, P. L., Drummond, C. K., Desai, M. Y., & Gillinov, A. M. (2020). Accuracy of the Apple Watch 4 to measure heart rate in patients with atrial fibrillation. *IEEE Journal of Translational Engineering in Health and Medicine*, 8, 1–4. <https://doi.org/10.1109/jtehm.2019.2950397>
 30. Singh, P. K., Kourav, P. S., Mohapatra, S. S. D., Kumar, V., & Panda, S. (2024). Human heart health prediction using GAIT parameters and machine learning model. *Biomedical Signal Processing and Control*, 88, 105696. <https://doi.org/10.1016/j.bspc.2023.105696>
 31. Stehlik, J., Schmalfuss, C., Bozkurt, B., Nativi-Nicolau, J., Wohlfahrt, P., Wegerich, S., Rose, K. A., Ray, R., Schofield, R. S., Deswal, A., Sekaric, J., Anand, S., Richards, D., Hanson, H., Pipke, M., & Pham, M. (2020). Continuous wearable monitoring analytics predict heart failure hospitalization. *Circulation: Heart Failure*, 13(3). <https://doi.org/10.1161/circheartfailure.119.006513>
 32. Wang, W., Brinker, A. D., Stuijk, S., & De Haan, G. (2017). Algorithmic principles of remote PPG. *IEEE Transactions on Biomedical Engineering*, 64(7), 1479–1491. <https://doi.org/10.1109/tbme.2016.2609282>
 33. Wijaya, R., Prihatmanto, A. S., & Kuspriyanto. (2013). Preliminary design of estimation heart disease by using machine learning ANN within one year. *Joint International Conference on Rural Information & Communication Technology and Electric-Vehicle Technology (rICT & ICeV-T)*, Bandung-Bali, Indonesia. <https://doi.org/10.1109/rict-icevt.2013.6741541>
 34. Wu, J., Kors, J. A., Rijnbeek, P. R., Van Herpen, G., Lu, Z., & Xu, C. (2003). Normal limits of the electrocardiogram in Chinese subjects. *International Journal of Cardiology*, 87(1), 37–51. [https://doi.org/10.1016/s0167-5273\(02\)00248-6](https://doi.org/10.1016/s0167-5273(02)00248-6)
 - Yu, J., Park, S. J., Kwon, S., Cho, K., & Lee, H. (2022). AI-based stroke disease prediction system using ECG and PPG bio-signals. *IEEE Access*, 10, 43623–43638. <https://doi.org/10.1109/access.2022.3169284>
 35. Zaman, R., Cho, C. H., Hartmann-Vaccarezza, K., Phan, T. N., Yoon, G., & Chong, J. W. (2017). Novel fingertip image-based heart rate detection methods for a smartphone. *Sensors*, 17(2), 358. <https://doi.org/10.3390/s17020358>
 36. Zhang, H., Zhang, P., Lin, F., Chao, L., Wang, Z., Ma, F., & Li, Q. (2024). Co-learning-assisted progressive dense fusion network for cardiovascular disease detection using ECG and PCG signals. *Expert Systems With Applications*, 238, 122144. <https://doi.org/10.1016/j.eswa.2023.122144>
 37. Zishan, M. O., Shihab, H., Islam, S. S., Riya, M. A., Rahman, G. M., & Noor, J.

(2024). Dense neural network based arrhythmia classification on low-cost and low-compute micro-controller. *Expert Systems With Applications*, 239, 122560. <https://doi.org/10.1016/j.eswa.2023.122560>