Research Article

International Journal of Advanced Technology and Engineering Exploration, Vol 6(54)
ISSN (Print): 2394-5443 ISSN (Online): 2394-7454
http://dx.doi.org/10.19101/IJATEE.2019.650027

An IoT framework for bio-medical sensor data acquisition and machine learning for early detection

Ayaskanta Mishra[1*] and Manaswini Mohapatro[2]
Assistant Professor, School of Electronics Engineering, Kalinga Institute of Industrial Technology (KIIT), Deemed to be University, Bhubaneswar, India[1]
Technology Analyst, Infosys Technologies Ltd., Info-city (Phase-I), Bhubaneswar, India[2]

Received: 11-March-2019; Revised: 24-May-2019; Accepted: 27-May-2019

Abstract
Internet of things in clinical domain has opened up new possibilities in remote monitoring of patients by connecting healthcare bio-sensor systems over the internet. This paper has proposed a working prototype of a real-time health monitoring system, which collects sensor data from body area network and communicates the data to a predictive model that is trained on historical clinical data. The prototype is equipped with Analog Device™ AD 8232 module for electrocardiogram and heart rate monitoring. CYPRESS CY8CKIT-042-BLE-A PSoC® 4 Bluetooth® Low Energy Pioneer Kit is used for implementation of a body area network, which collects patient's vitals and communicates the sensor data to a Raspberry Pi3. The gateway device between WPAN (Bluetooth® Low Energy) and WLAN (IEEE 802.11n) is implemented using Raspberry Pi3. The gateway device collects the sensor data over a Bluetooth personal area network coming from all the connected devices and the data is acquired over internet server. ECG- ST wave and heart rate data are sent to the cloud server from the sensors. On the server, a machine learning model is deployed to predict any malfunctions based on sensor readings posted from the real-time health monitoring system and generate early alerts. We have obtained >90% prediction accuracy with random forest classifier using the UCI heart diseases repository.

Keywords
Internet of things, Machine learning, Body area network, Analog deviceTM AD 8232, Electrocardiogram, CYPRESS CY8CKIT-042-BLE-A PSoC® 4 Bluetooth®, Raspberry Pi3, Cloud server.

1.Introduction
Technology is an indispensable part of our day-to-day activities and major driving force behind revolutionizing every aspect of our lives by providing automated solutions and simplifying various processes. New innovations are no longer limited to providing services and solutions, rather technological advances in IoT and machine learning are making it possible for smart devices collect real time data and translate that data to understand the greater needs of the society, thus being the harbinger of artificial intelligence.

*Author for correspondence

This work was supported by the School of Electronics Engineering, KIIT Deemed to be University, India under Cypress Semiconductor Academic Alliance (PSoC4 BLE kit is provided by Cypress Semiconductors).

Through big data, cloud computing and machine learning algorithms, the data obtained over a period is trained to develop predictive models that can help to filter and prioritize the vital parameters with or without human supervision.

Patients' health care is a very intensive and intricate process and cardiac arrest that affect a huge population often remain undetected and lead to fatal conditions. Tests such as coronary angiography, MRI scans, CT Scans can detect coronary heart diseases such as cardiomyopathy, clogged arteries, etc., but physicians cannot predict the exact time of occurrence of cardiac arrest.

The aim of the paper is to develop a heterogeneous IoT system integrated with machine learning for developing an intelligent alarming system for efficient risk prediction of cardiac arrest. With the help of connected smart devices the patient can be

under constant remote monitoring while performing normal daily activities.
The objectives are outlined as below:

- Design a heterogeneous IoT body area network (BAN) and communicate ECG and heart bps to the server.
- Real time processing of ECG and classify irregularities in the wave.
- Develop a machine learning model using Random Forest Classifier.
- Real time prediction of the data acquired.

In short, the paper focuses on developing a model for real-time health monitoring system (RTHMS) for remote bio-sensor data acquisition from smart wearable and communicating it over the internet to a predictive model for identifying coronary complications that can generate early alert for preventive measures.

With the growing number of populations, traditional health care systems are fast becoming obsolete, giving way to IoT based pervasive healthcare systems. A framework of the health care system using IoT has been proposed in [1]. In their analysis they have suggested a LR-WAN network for fetching data from sensors and sending data to a machine learning based cloud. Results obtained from [1] suggest that BLE is more preferable option over Zigbee as a WBAN technology. In [2] the authors have developed a prototype for continuous monitoring of patient's physical signs like blood pressure, ECG, SpO2. In their proposed model they have taken four models to specify the sensor data traffic and how data should be sent to the cloud. Continuously, special period, Event triggered and patient demand. They have used a PDA running on android platform based on Java to acquire data using Bluetooth. In [3] the authors have suggested Body area network architecture for remote medical monitoring. They have suggested a 3-tier system with discrete sensors, data hub and medical networks. They have suggested BLE for sensor and data hub communication and GSM and other network to connect from data hub to medical network [4]. The authors have developed a smart phone-based platform for real-time detection of cardiovascular disease using ECG processing. They have used the Alive Technology ECG sensor to collect 300 8-bit samples per seconds over a class-1Bluetooth link to the Windows based PDA. They have used feature extraction and classification of artificial neural network and is implemented in the Smartphone itself and SD card is used for sensor data storage. In [5] the authors have proposed a FIR filter for de-noising and extraction of premature ventricular contraption (PVC) event from the ECG signal using AD8232 Module and proposed a sequential recursive (SR) Algorithm for data processing. They have designed an IoT expert system using STM32F4 microcontroller for data collection. The authors have done a study of various machine learning techniques for disease classification and prediction based on heart failure subtypes [6]. Ideas from various other research papers are integrated in material and method section.

2.Materials and methods

2.1IoT-RTHMS framework

Data acquisition from a wearable device for real-time monitoring of patient vitals is a key to this work. The scope of this work includes a proposed system for bio-sensor data acquisition and communicating the same to a cloud server [7]. This can be achieved using a heterogeneous network architecture based on IoT for real-time sensor data acquisition with last mile as Bluetooth 4.0 to harness the advantage of Bluetooth low energy (BLE) as proposed in [8]. BLE is particularly chosen in this application keeping very low power coin cell battery powered wearable embedded device in mind. *Figure 1* shows the proposed system for collecting data from bio sensor. Bio sensor is to be integrated to a low-power microcontroller (MCU) based embedded device [9].

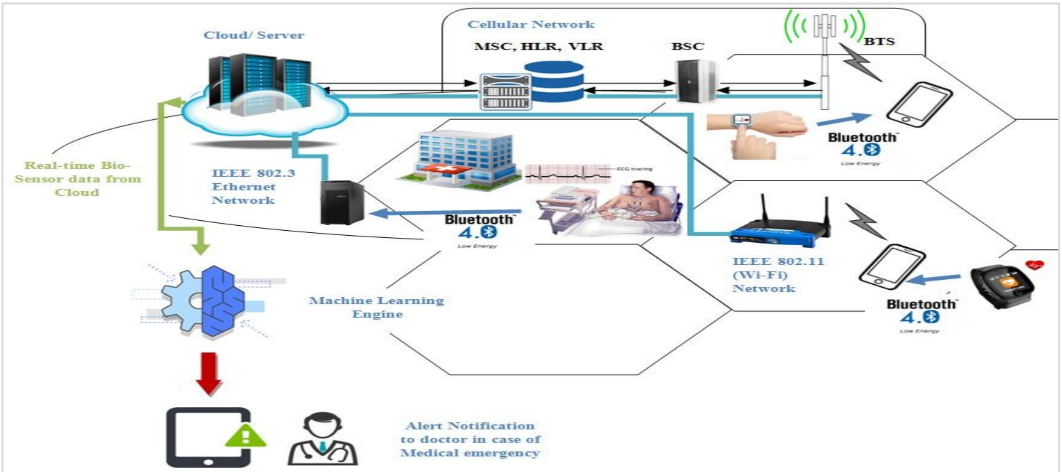


Figure 1 Proposed heterogeneous IoT network architecture for RTHMS integrated with machine learning modalities

2.2IoT-RTHMS system design

Modern days Bio Sensors are single chip Integrated circuits designed to give the various parameters like subject's heart-rate, blood pressure, blood sugar level [10]. In this work we have proposed a cypress semiconductor PSoC4 BLE based prototyping board. For proof of concept we have used a CY8CKIT-044 PSoC® 4 M-Series pioneer kit. Though the firm factor of this prototyping board is not small enough for wearable devices [11], this board is used for only prototyping and wearable products can be developed using cypress PSoC 4 IC with custom board with smaller firm factor by implementing similar algorithm and software. The last mile of this heterogeneous network is BLE, however the backbone of this proposed network architecture may support heterogeneous technology [12] based on the customer requirements [13]. The idea proposed here is to implement a heterogeneous wireless network to send the Cardiac Monitor module data to the cloud server for analytics [14]. *Figure 2* depicts the proposed system block diagram to connect the AD8232 low power cardiac monitor module to the cloud server. The idea is to design a biomedical sensor network using IoT, which is similar to some other work in this field [15, 16].

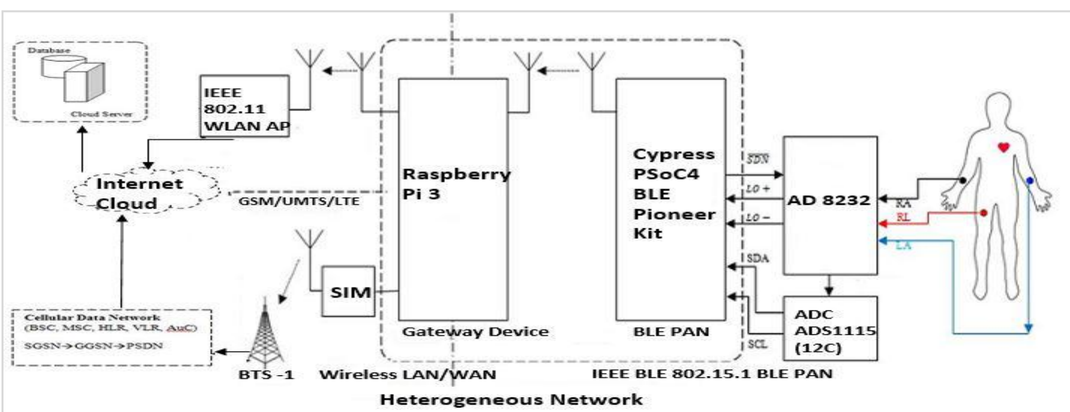


Figure 2 Block diagram of heterogeneous network framework implementation for IoT-RTHMS

In this prototype ADS1115, 10-bit ADC is used to convert the ECG analog signal to a 10-bit digital. I2C interface is used to interface the ADC with the Cypress PSoC4 pioneer kit using data SDA and clock SCL pins. The ADS1115 is a multi-channel ADC out of which A0, only one channel is used as in this application only one analog output from AD8232 is fed to the PSoC4 MCU. After the cardiac sensor data is collected from I2C interface the PSoC4 sends the data using a BLE module to the gateway device (Raspberry Pi). Bluetooth low energy 4.0 is used particularly by keeping a low power application in mind. The BLE module is configured as a peripheral device in the GATT service. The point here is to pair the PSoC4 BLE device to the Raspberry Pi 3, which is working as a gateway between the Bluetooth personal area network (PAN) and a WLAN/WAN. Here in this work we have also proposed a multi technology platform either WLAN or Cellular WAN to connect the gateway device to the internet cloud. *Figure 3* shows the picture of our implemented hardware to use IEEE 802.11 Wireless Local Area Network (WLAN) as the backbone to connect the BLE to cloud through the Raspberry Pi 3 as the gateway. The Bluetooth PAN side of the gateway (Raspberry Pi) is used for acquisition of cardiac sensor data using BLE GATT service form PSoC4. At the PSoC4 device the Analog Device's AD8232 cardiac sensor is connected through a 10-bit ADC uses I2C interface. AD8232 is giving the output in analog voltage in the range 0 - 3.3V (Vcc).

The ADC is programmed to get the ECG data in required samples per second (SPS). The gateway device, Raspberry pi is paired with PSoC4 BLE. The PSoC4 is working as a peripheral device and constantly sending the cardiac sensor data to the Raspberry Pi over a Bluetooth PAN. The Raspberry Pi stores the sensor data in a file for 1-minute duration. The value of individual reading (sample) can be "0-1023" (10bit ADC) for each sample of analog data with sampling frequency 50 Hz as presented by *Mahdiani et al*. in their research findings [17]. The Texas instrument's ADS1115 ADC is programmed to 50 SPS. The heart rate in beats per minute (BPM) is calculated locally from the 1-minute data. The BLE device sends 50 SPS x 60 seconds = 3000 samples/ minute. The data-rate requirement of this application on BLE network can be given as 10-bit x 50 SPS = 500 bps. After receiving the data at the Raspberry Pi, these 3000 samples are stored in a JSON file format with the time stamp. The proposed framework implements an edge computing technique to find out three slope characteristics of the ST wave from the JSON file every minute at the Raspberry Pi. This saves a lot of computational resources in the cloud. Further, this technique helps reducing the data required to be sent to cloud periodically (each minute). Instead of sending all raw sensor data (3000 samples/minute) this proposed edge computing technique only sends one of three numerical values to the internet cloud from the gateway device over IEEE 802.11 network using REST API (HTTP POST). Only ST wave slope characteristic: *1: up sloping; 2: flat; 3: down sloping and Heart rate (in BPM)* are sent to the cloud server. The raspberry pi is having a python module to read the JSON file raw sensor data and find out the slope of an ECG ST wave and give 1, 2 or 3 as output. The Raspberry Pi is connected to an Access point, which connects to the ISP so that it can be seamlessly connected to the internet. For a typical implementation internet socket () program is running on the Raspberry Pi to connect to the cloud server, which is deployed on a particular URL. ST wave characteristics and heart rate, which are calculated every minute from the data received from the BLE PAN is sent to the cloud URL uses internet socket () program. *Figure 3* shows the implementation of heterogeneous networking using BLE and IEEE 802.11 WLAN via a Raspberry Pi gateway device.

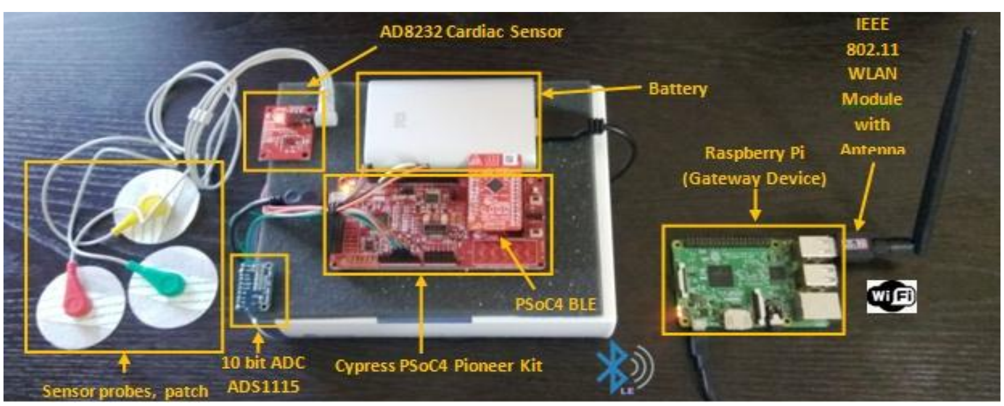


Figure 3 AD8232 low power cardiac monitor cypress PSoC4 BLE connected to Raspberry Pi3 wireless LAN gateway device connected to cloud server

Figure 4 shows the implementation of heterogeneous network using a GSM cellular WAN. This is an alternative of the IEEE 802.11 WLAN. As discussed above, we acquire the cardiac sensor data using the

BLE PAN from PSoC4 to the gateway device (Raspberry Pi). Now, instead of the WLAN a Cellular data network is used for connecting to the internet through PSDN. The SIM900 is interfaced with the Raspberry Pi gateway device using the UART interface. The python program running on the Raspberry pi can send the AT command to SIM900 module for an instant "AT+CIPSEND" command to connect to the cloud server using the HTTP 1.1 POST message over a TCP connection.

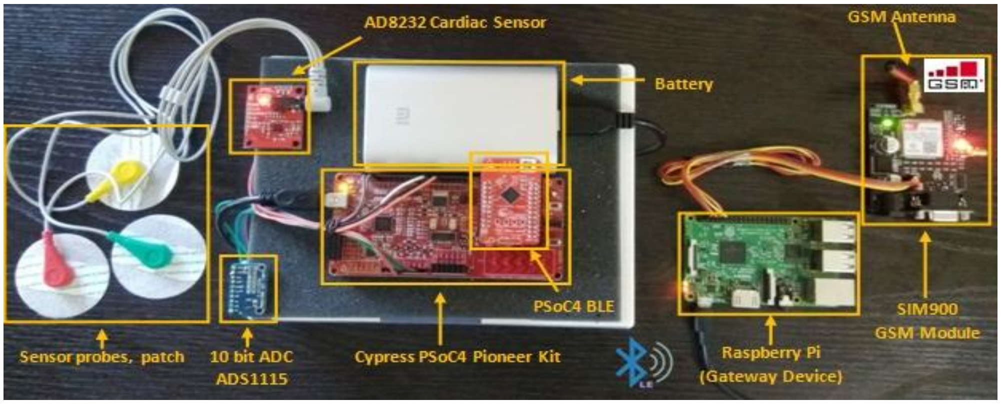


Figure 4 AD8232 low power cardiac monitor cypress PSoC4 BLE connected to Raspberry Pi3 SIM900 cellular network PSDN WAN gateway device connected to cloud server

The cloud server is deployed using an HTTP, PHP with MySQL database. In case of first implementation using IEEE 802.11 network the socket program running on Raspberry Pi will send the data to the cloud server using HTTP POST. The data is captured using a PHP script running on the server and capture the data and store in a MySQL database using a database handler. After fetching ST wave slope characteristics and heart rate (BPS) data to the MySQL database with the appropriate timestamp an analytics and predictive machine learning algorithms can be implemented in the cloud for early detection of cardiac anomaly for the patient.

The following datasets are taken from the MySQL for consideration of the parameters for cardiac disease.

Parameters for data analytics:

Type of irregularities based on ST wave of ECG:
1: up sloping; 2: flat; 3: down sloping and Heart rate in BPS

The machine learning model will take these parameters as input for predicting likelihood of coronary complications.

2.3Machine learning model

The data considered for training the predictive model will focus on the parameters that lead to coronary complications and the model is trained in a way that the real time values of these parameters can be remotely monitored to predict future incidents of coronary failures.

For building machine learning models, the Python programming language is widely used as it has very powerful data science libraries. The prerequisites for building a predictive model are discussed in [18, 19]. Open source Python distribution package Anaconda consists of various IDEs such as Sypder and Jupyter IPython Notebook that are used in our model for building and testing the machine learning codes. *Figure 5* represents the flow diagram for developing a machine learning model.

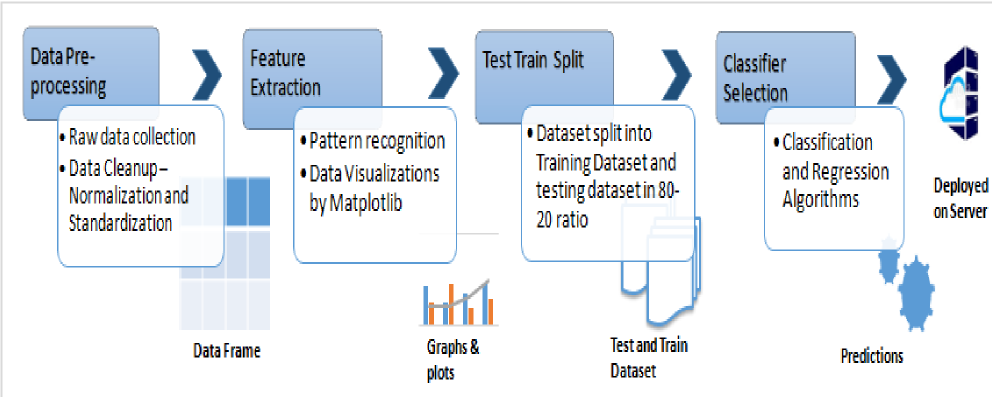


Figure 5 Flow diagram for developing a machine learning model

2.3.1Data pre-processing

The first step for building a predictive model is a collection of raw data from various resources such as past hospital records, doctor prognosis reports, test reports. For the prototype model in this paper we have considered data from the UCI's machine learning repository, this data can be found on the UCI's machine learning repository [20, 21]. It consists of 270 samples with 13 attributes each. Machine learning researchers have used this database to predict presence of heart disease in patients. We have used 11 attributes from the set of 13 attributes for building the proposed machine learning model. *Table 1* shows the list of attributes and their equivalent mapped values from the raw data.

Table 1 List of attributes and equivalent mapped values from raw data collected

Attributes	Mapped Value/Unit
Patient ID	-
Age	-
Gender	1 - Male; 0 -Female
Resting Blood Pressure	-
Cholesterol	in mg/dl
Fasting Blood Sugar (in >120 mg/dl)	1 - true; 0 – false
Rest ECG	0, 1, 2
Thalach: maximum heart rate	BPM
Exercise induced angina	1 - true; 0 - false
Slope of the peak ST segment	1: up sloping; 2: flat; 3: down sloping
Diagnosis of Coronary complication	1-<50% possibility; 2- ≥ 50% possibility

The normal resting blood pressure is 120/80, the attribute values in the table are the recorded systolic blood pressure of the patients. Cholesterol is measured in mg/dL and normal limit is 200 mg/dl. Fasting blood sugar is recorded blood sugar of patients during a checkup. Exercise induced angina is when the oxygen supply to heart decreases and the patient starts having trouble breathing. The heart beats per second is the maximum heart beat obtained during exercise. The slope ST wave is calculated from ECG wave. The diagnosis of heart disease is in binary format, either 1 - > 50% chance of coronary complication or 0 - <50% chance of coronary complication. *Table 2* shows sample attribute values from the UCI's machine learning repository saved to an excel sheet.

Table 2 Sample data sheet from UCI's machine learning repository with 11 attributes and attribute values

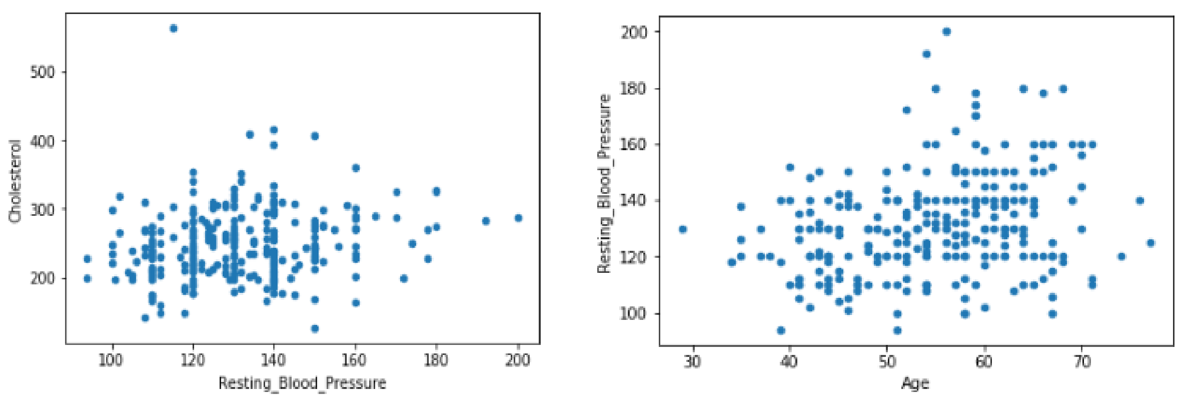
Age	Sex	Resting blood pressure	Cholesterol	Fasting blood sugar	Rest ECG	Max heart rate	Exercise induced angina	Slope ST wave	Diagnosis
70	1	130	322	0	2	109	0	2	2
67	0	115	564	0	2	160	0	2	1
57	1	124	261	0	0	141	0	1	2

Age	Sex	Resting blood pressure	Cholesterol	Fasting blood sugar	Rest ECG	Max heart rate	Exercise induced angina	Slope ST wave	Diagnosis
64	1	128	263	0	0	105	1	2	1

The excel sheet is converted to data frame via a numerical Python module (Numpy) and Python data analysis library (Data Panda) by executing the function pd.read_excel function. Data frames (df) are multi-dimensional arrays for computation of data. Before feeding the dataset to train the classifier model datasheet is pre-processed as in [22] by validating the sheet for missing values or inconsistent information. The missing attribute values are either being replaced with a mean value via normalization and standardization or deleted from the dataset.

2.3.2Feature extraction

To train the machine learning model we need to analyze the correlation between the data attributes and how they affect the diagnosis of the disease. Pattern recognition and feature extraction, extract the determining factors from huge datasets and render information on primary attributes that lead to coronary conditions. This is a significant step in developing the machine learning model as in [23]; the model is trained on the primary attributes. To analyze the data from the datasets, we have employed Matplotlib library from Python data science which has built-in functions such as plot_corr(), hist(), pd.crosstab() that can be used for data visualization. *Figure 10* represents the correlation between various attributes via scatter plots, *Figure 6(a)* is the scatter plot for cholesterol and blood pressure showing maximum values for cholesterol>200mg/dl scattered when resting blood pressure>120. *Figure 6 (b)* is the scatter plot for age and resting blood pressure showing age being related to blood pressure.



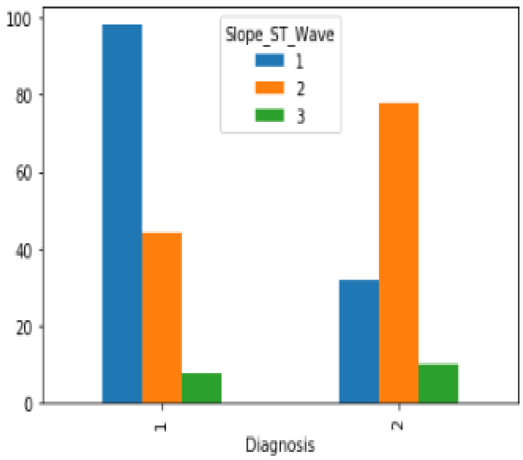
(a) (b)

Figure 6 Scatter Plots for visualizing the correlation between various attributes

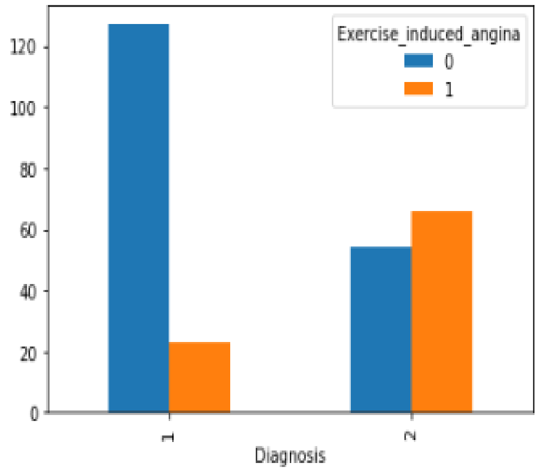
Figure 7 shows the frequency density of patients diagnosed with heart disease with (a) ST wave slope of ECG and (b) Exercise induced angina. The features that are more likely to affect the diagnosis are taken into the dataset. The dataset obtained after feature extraction is split into two subsets–train dataset and test dataset. The training dataset is applied to learn the pattern and to validate the correlation of the test dataset with model predictions. Train_test_split function is imported from a data science library that splits the dataset with the mentioned test and train sizes.

2.3.3Classifier selection

There are various supervised and unsupervised machine learning algorithms that can be applied on the training dataset. Supervised machine learning algorithms can be classified based on the logic it employs which can be either classification or regression. Classification algorithms are applied when output is from a finite set of outcomes. Multi-label classification algorithms in machine learning are decision trees, logistic regression and K-nearest neighbor (KNN). Regression is used for predictions those are quantifiable and not confined to probable binary outcomes. Decision Trees are the simplest classifier models using both classification and regression for its predictive algorithm. As discussed in [24], decision tree outperforms other machine learning techniques (artificial neural network, naive Bayes) with greater prediction accuracy. The reason behind choosing random forest classifier for our work is because this algorithm is based on multiple decision trees and employs both classification and regression for its predictive algorithm. Random forest algorithm is described as a collection of randomized base regression trees that are combined together to give an average regression estimate [25]. Although our output is in binary results and a classification type of algorithm would be more suitable, but for future work we have used a regression type of algorithm.



(a)



(b)

Figure 7 Bar Graphs for visualizing frequency density of patients diagnosed with heart disease with (a) ST wave slope and (b) exercise induced angina

In [26], the authors have presented a detailed illustration of attribute selection measures for the random forest classifier. Another reason for using random forest classifier is it reduces the variance as it is trained on different samples of the data and takes random subset of features for all iteration. The prediction accuracy of the above algorithm on our dataset is 99%. *Figure 8* is the confusion matrix showing 98% of values were correctly diagnosed as 1 i.e. the true values is same as the predicted values.

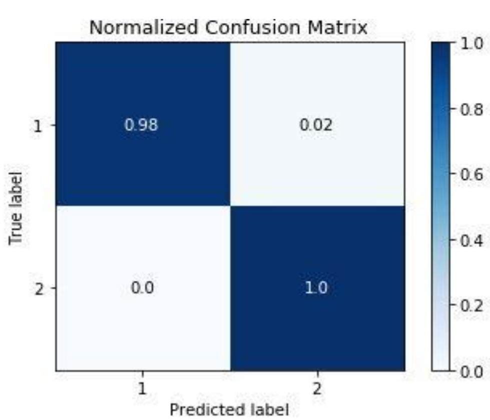


Figure 8 Confusion matrix for predicted and true values from the test dataset

2.3.4 Deployment

For deployment on the server, the model is converted to a serialized file by a process called pickling which requires Python object serialization library. The pickle file (.pkl) generated is integrated in a Django app hosted on the cloud server.

The heterogeneous IoT network in section 2.2 sends ECG and heart bps to the Django app. An ECG can be any one of 1, 2 or 3 based on the ST wave. This data along with individual's other vitals from medical test reports are fed to the machine learning model every minute for continuous monitoring. In case result of the app being 2, data is communicated to care taker via an SMS feed.

2.4Algorithm for RTHMS software development

The hardware design requires a particular software implementation for development of the working prototype for the biomedical sensor data acquisition system. For implementation of this RTHMS we have developed the software using embedded C for PSoC4 BLE, python for Raspberry Pi and PHP and Java Script with MySQL database for cloud Server respectively.

2.4.1Description of subroutines for sensor nodes

The PSoC4 BLE is used for prototyping of the Cardiac Sensor data acquisition module deployed in

Bluetooth PAN Network. Embedded C language is used in programming the Cypress PSoC4 BLE device. The AD8232 cardiac sensor data acquisition is done using I^2C interfacing of AD1115 ADC module. The main program flow chart of processes running on PSoC4 BLE is shown in primarily the system analog sensor data acquisition through an I^2C interface and the sensor data is sent to the Raspberry Pi gateway device using a BLE WPAN.

2.4.2Description of subroutines for gateway device

Figure 9 shows the flow chart of processes running in Raspberry Pi, the gateway device. Python programming is used to develop the software, which can fetch the data from BLE network and forward it to IEEE 802.11 or a cellular network using SIM900. This flowchart shows the software implementation of deployment of a heterogeneous wireless network (HetNet) particularly for a Bio medical application. The gateway device designed in this context is working on the concept of HetNet. This system will leverage the advantage of WPAN is access network and WLAN in the backbone. This scenario is best fitted for sensor network architecture especially for bio-medical application. Our proposed hybrid architecture support both IEEE 802.11 WLAN as well as a GSM based data network on the backbone side. This is more suitable for any practical deployment scenario for urban as well as rural scenario with multiple network technology. Hence, proposed system will not be constrained by network availability and will support multiple technologies.

2.4.3Description of subroutines for cloud server

Figure 10 shows the flow chart of processes running on the cloud server. The data is captured by the PHP Server using RESTful API HTTP POST messages. The captured Biomedical Sensor data are then stored in MySQL database using database handler.

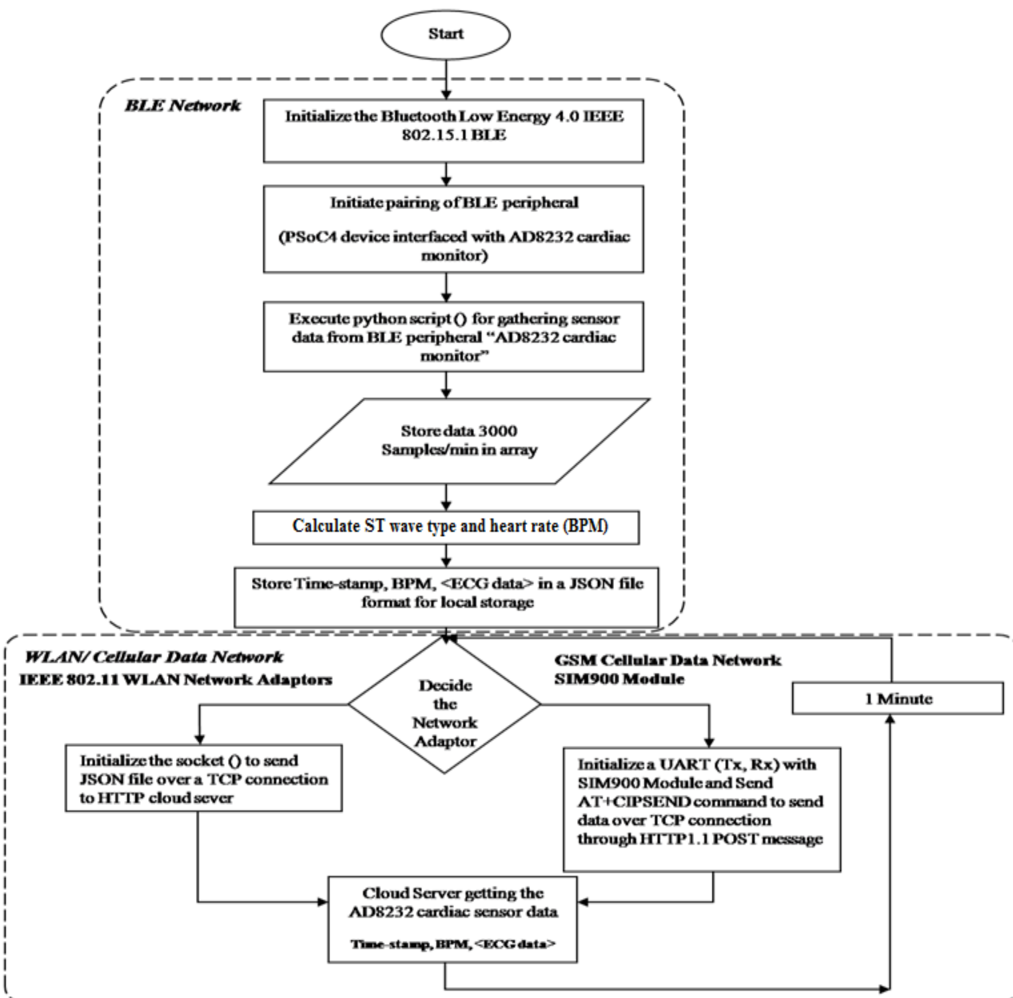


Figure 9 Flow Chart of the processes running on the Raspberry Pi, Gateway device

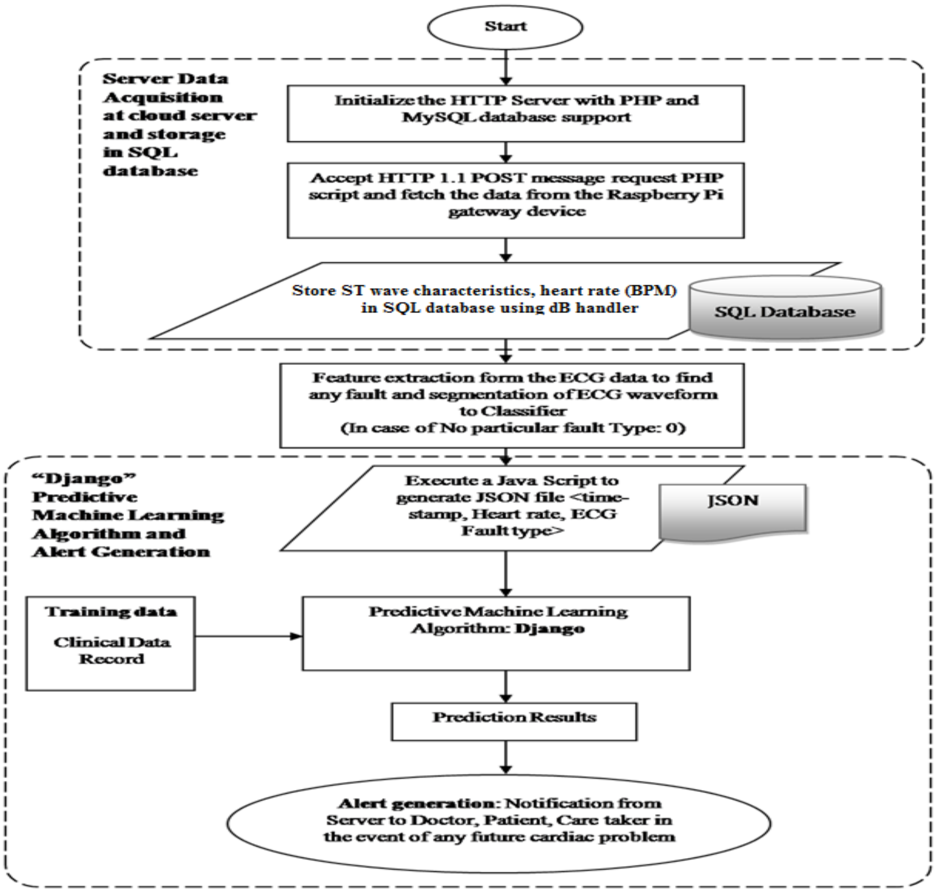


Figure 10 Flow chart of the processes running on the cloud server

3.Results

Figure 11 shows a sample ECG sensor data obtained from AD8232 sensor using the PSoC4 BLE at Raspberry Pi. This data is stored in a JSON file at the gateway device for determining the ST wave slope characteristics using the python module as discussion section 2.2. *Figure 12* shows the three ST wave slope types for determining cardiac anomalies using the machine learning model as discussed in section 2.3.

From the ST wave characteristic 1, 2 or 3 values along with heart rate (BPM) is sent to the internet cloud of storage. ST wave and heart rate data is stored in MySQL database using PHP and dB handler. This data is used as the input to machine learning model. *Figure 13* shows the MySQL visualization tool for data acquired in cloud server. *Figure 14* showing heart rate data in cloud.

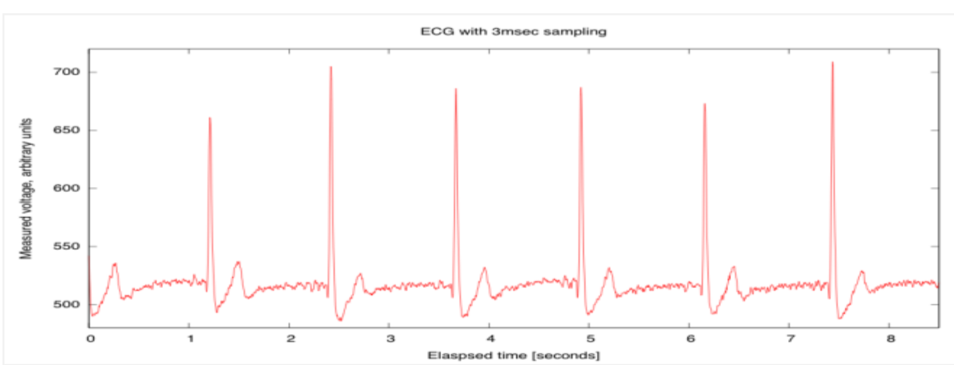


Figure 11 Visualization of JSON file showing ECG collected from the AD8232 sensor at Raspberry Pi

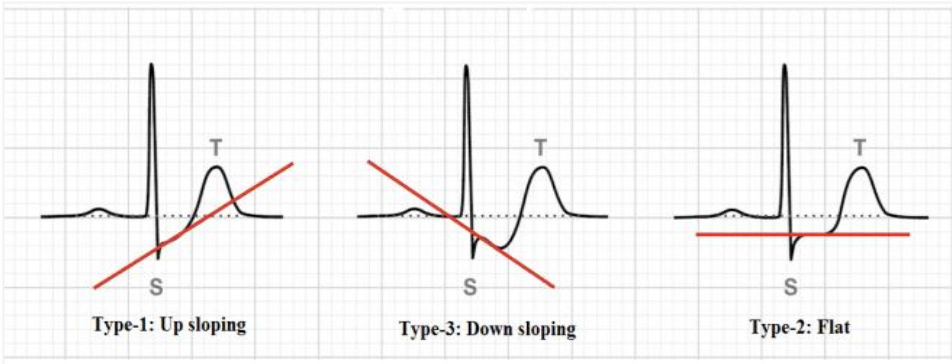


Figure 12 Determining the ST wave slope characteristic types for input the machine learning model

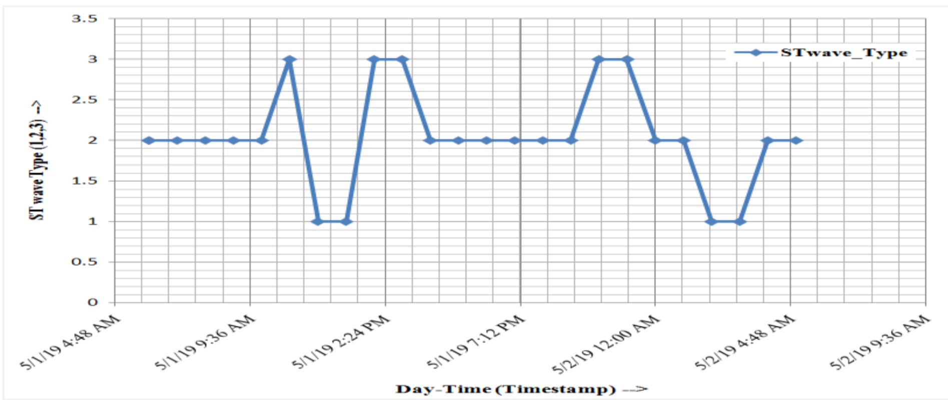


Figure 13 Visualization of ST wave type 1-up sloping, 2- flat or 3- down sloping at cloud server

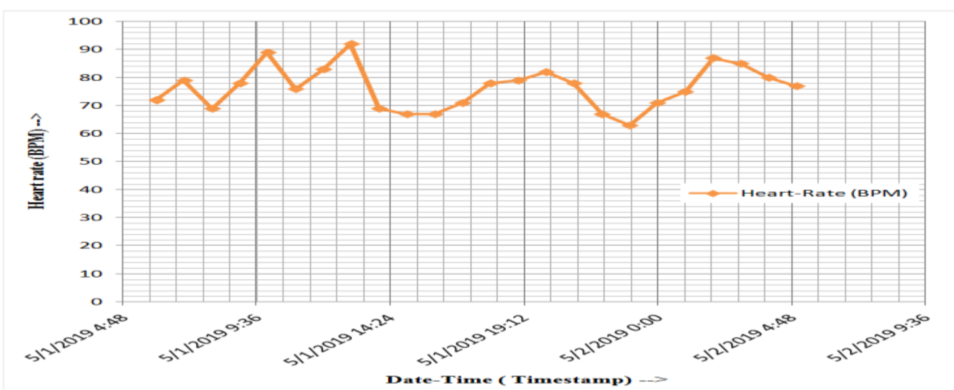


Figure 14 Visualization of ST wave type 1-up sloping, 2- flat or 3- down sloping at cloud server

The ST wave type and heart rate data are acquired at cloud server using the proposed heterogeneous network i.e. BLE in PAN and WLAN using gateway device as discussed in section 2.2. The detailed algorithm and software subroutine have been discussed in section 2.4. *Figure 15* shows the real time sensor data in MySQL database managed by phpMyAdmin and *Figure 16* shows the sensor data accessed from the cloud using an Android Smartphone using standard HTTP running on PHP.

As discussed in section 2.3 a machine learning model in deployed using the testing data as shown in the above results. The above result shows the acquisition of real-time ECG (ST wave) and heart rate sensor data to the cloud. These results are the input to a Django app running on the server as {ST wave, heart rate} for predicting any possibilities of cardiac complexities. The response from the Django app is a prediction quantifier. If the possibility of cardiac arrest is ≥50%, then the output value is 2 and in case of < 50% the output is 1. We have obtained >90% prediction accuracy with Random Forest classifier using the UCI heart diseases repository.

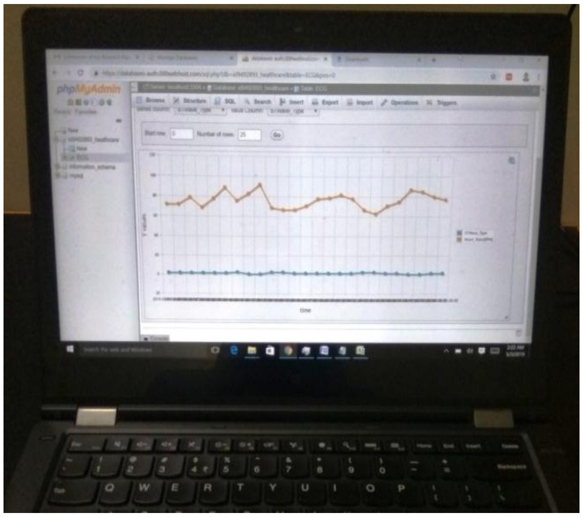


Figure 15 Photo of phpMyAdmin for SQL database management deployed in the cloud server for sensor data

4.Discussion

Table 3 shows comparison with other related work [2, 4, 5, 15] similar research where the ECG is acquired over an IoT Network.

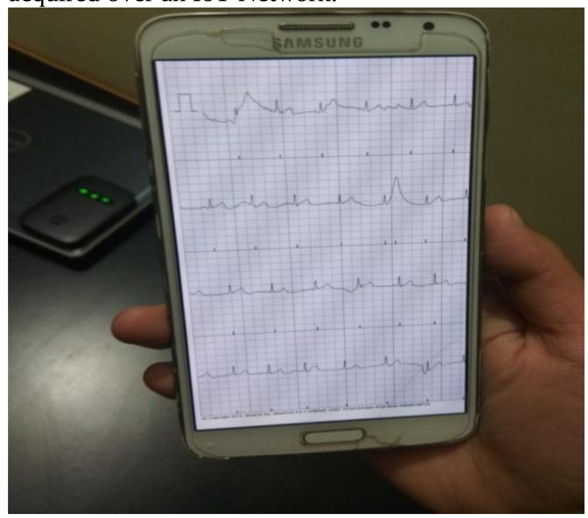


Figure 16 The ECG sensor data on an Android Smartphone

Table 3 Comparison with related work in IoT for healthcare

Features	[2]	[4]	[5]	[15]	This work
Embedded Platform	Android based device	HeartToGo prototype based on Windows Mobile platform	STM32F4	ATmega328P	Cypress PSoC4 BLE Edge device and Raspberry Pi Gateway Device
Sensors	Not specified	Alive Technology ECG Sensor	AD8232	Accelerometer-ADXL362, Temperature Sensor MAX30205, Pulse Sensor	AD8232
Programming Language	Java	C#	Embedded C	Embedded C	Embedded C PSoC, python for Raspberry pi and html java, php for web GUI
Communication Technology	Bluetooth	Bluetooth Serial port	Bluetooth	BLE-Texas Instrument's CC2541	Adaptive/ on demand BLE PAN with Gateway with 4G/5G WLAN support
Heterogeneous Architecture	No	No	No	No	On demand IEEE 802.11/ 4G/5G support
Power	Higher than BLE	Higher than BLE	Higher than BLE	Low Power	Low power with additional features: Continuous/Scheduled/ Event driven case specific duty cycle for adaptive power requirements
Cloud Technology	Not implemented	NA	NA	NA	Own cloud based on HTTP, PHP, MySQL Sensor data acquisition using RESTful API
Database	Local storage on Mobile device	Local storage on Mobile device	Not specified	Local storage on Mobile device	MySQL, Global support for different data types

Features	[2]	[4]	[5]	[15]	This work
User Interface	Android/ Doctors GUI	Windows Mobile App	Not specified	Smart phone	Any platform, Android, iphone, Windows, Linux, WebUI
Learning Model	N.A	ANN	Sequential Recursive Algorithm	N.A	Random Forest
Training database and Size	N.A	MIT database	MIT-BIH	N.A	UCI Medical database (x MB)
Processing Nodes	N.A	Edge Node only	Edge node only	N.A	Sensor data pre-processing at Edge device, Machine Learning at Cloud to optimize power
Processing time x power	Medium	Medium	Medium	N.A	() PSoC4 BLE + () Gateway device
Result	No prediction	81-98 % for different abnormalities	PVC incident detection	N.A	99 % accuracy from MACHINE LEARNING model

4. Conclusions

With the help of connected smart devices proposed as RTHMS, patient with coronary conditions can be under constant remote monitoring while performing normal daily activities. Results obtained from the RTHMS prototype show that remote acquisition of ECG and heart bps along with clinical parameters cholesterol, blood sugar can be evaluated by a machine learning model to predict whether the vitals might lead to cardiac arrest. This work can give a direction on practicable implementation of machine learning used for patient risk analysis remotely.

Challenges that can be addressed to design a more scalable system by further extending this work are:

- Using various IoT enabled devices such as smart gluco-meter, bio-patch to provide vital data from sensors for constantly monitoring the patient without interfering with day-to-day activities.
- From the UCI repository for diagnosis of heart health we have considered only 9 parameters. More parameters can be trained in the machine learning model for accurate predictions.
- We have limited our work to extract features from the ST wave of the ECG. Research on R peak of the ECG wave can lead to precise results.

In the future, we plan to design a RTHMS that works seamlessly and provide multi-functional prediction mechanism for other health conditions.

Acknowledgment

The prototypes are developed in Wireless Communication & Networking Lab, School of Electronics Engineering, Kalinga Institute of Industrial Technology (KIIT) Deemed to be University. All the hardware and software vendors are acknowledged and no commercial grade or licensed software is used for this work. This work is for purely academics and research.

Conflicts of interest

The authors have no conflicts of interest to declare.

References

[1] Baker SB, Xiang W, Atkinson I. Internet of things for smart healthcare: technologies, challenges, and opportunities. IEEE Access. 2017; 5:26521-44.

[2] Li C, Hu X, Zhang L. The IoT-based heart disease monitoring system for pervasive healthcare service. Procedia Computer Science. 2017; 112:2328-34.

[3] Jurik AD, Weaver AC. Remote medical monitoring. Computer. 2008; 41(4):96-9.

[4] Oresko JJ, Jin Z, Cheng J, Huang S, Sun Y, Duschl H, et al. A wearable smartphone-based platform for real-time cardiovascular disease detection via electrocardiogram processing. IEEE Transactions on Information Technology in Biomedicine. 2010; 14(3):734-40.

[5] Thai DT, Minh QT, Phung PH. Toward an IoT-based expert system for heart disease diagnosis. CEUR Workshop Proceedings. 2017:157-164.

[6] Austin PC, Tu JV, Ho JE, Levy D, Lee DS. Using methods from the data-mining and machine-learning literature for disease classification and prediction: a case study examining classification of heart failure subtypes. Journal of Clinical Epidemiology. 2013; 66(4):398-407.

[7] Fiaidhi J, Mohammed S. Digital health in the era of extreme automation. IT Professional. 2018; 20(3):90-5.

[8] Hossain MS, Muhammad G. Emotion-aware connected healthcare big data towards 5G. IEEE Internet of Things Journal. 2018; 5(4):2399-406.

[9] Alam MM, Malik H, Khan MI, Pardy T, Kuusik A, Le Moullec Y. A survey on the roles of communication technologies in IoT-based personalized healthcare applications. IEEE Access. 2018; 6:36611-31.

[10] Chen X, Rhee W, Wang Z. Low power sensor design for IoT and mobile healthcare applications. China Communications. 2015; 12(5):42-54.

[11] Yang P, Stankevicius D, Marozas V, Deng Z, Liu E, Lukosevicius A, et al. Lifelogging data validation

model for internet of things enabled personalized healthcare. IEEE Transactions on Systems, Man, and Cybernetics: Systems. 2018; 48(1):50-64.
[12] Salahuddin MA, Al-Fuqaha A, Guizani M, Shuaib K, Sallabi F. Softwarization of internet of things infrastructure for secure and smart healthcare. Computer. 2018; 50(7):74-9.
[13] He S, Cheng B, Wang H, Huang Y, Chen J. Proactive personalized services through fog-cloud computing in large-scale IoT-based healthcare application. China Communications. 2017; 14(11):1-16.
[14] Muhammed T, Mehmood R, Albeshri A, Katib I. UbeHealth: a personalized ubiquitous cloud and edge-enabled networked healthcare system for smart cities. IEEE Access. 2018; 6:32258-85.
[15] Wu T, Wu F, Redouté JM, Yuce MR. An autonomous wireless body area network implementation towards IoT connected healthcare applications. IEEE Access. 2017; 5:11413-22.
[16] Guo J. Smartphone-powered electrochemical biosensing dongle for emerging medical IoTs application. IEEE Transactions on Industrial Informatics. 2018; 14(6):2592-7.
[17] Mahdiani S, Jeyhani V, Peltokangas M, Vehkaoja A. Is 50 Hz high enough ECG sampling frequency for accurate HRV analysis? In international conference of the engineering in medicine and biology society 2015 (pp. 5948-51). IEEE.
[18] Holzinger A. Machine learning for health informatics: state-of-the-art and future challenges. Springer; 2016.
[19] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. Journal of Machine Learning Research. 2011; 2825-30.
[20] http://archive.ics.uci.edu/ml/index.php. Accessed 25 November 2018.
[21] https://archive.ics.uci.edu/ml/datasets/Heart+Disease. Accessed 25 November 2018.
[22] Shahul S, Suneel S, Rahaman MA, Swathi JN. A study of data pre-processing techniques for machine learning algorithm to predict software effort estimation. Imperial Journal of Interdisciplinary Research. 2016; 2(6): 546-50.
[23] Aparna UR, Paul S. Feature selection and extraction in data mining. In online international conference on green engineering and technologies 2016 (pp. 1-3). IEEE.
[24] Soni J, Ansari U, Sharma D, Soni S. Predictive data mining for medical diagnosis: an overview of heart disease prediction. International Journal of Computer Applications. 2011; 17(8):43-8.
[25] Biau G. Analysis of a random forests model. Journal of Machine Learning Research. 2012; 13: 1063-95.
[26] Gupta B, Rawat A, Jain A, Arora A, Dhami N. Analysis of various decision tree algorithms for classification in data mining. International Journal of Computer Applications. 2017; 8:15-9.

Ayaskanta Mishra is working as Assistant Professor, School of Electronics Engineering, Kalinga Institute of Industrial Technology (KIIT), Deemed to be University, Bhubaneswar, India. He is currently heading the "Wireless Communication and Networking Laboratory" in his department. He has been a member of IET since 2017, a life member of the ISCA since 2012. He had received his Bachelor of Engineering degree from Sathyabama Institute of Science and Technology, Deemed to be University, Chennai in 2007. He worked as an Engineer at Siemens Communication and Enterprise Networks during 2007-08. He finished his Master of Technology from International Institute of Information Technology, Pune in the year 2011. His main research interest includes Internet of Things (IoT), Machine learning, Heterogenous Wireless Networks (HetNet), Cognitive Radio, Next-Generagtion-Networks (NGN). He has 7 years of teaching experience and 3 years of research experience.
Email: ayaskanta.mishrafet@kiit.ac.in

Manaswini Mohapatro is currently working as a Technology Analyst at R&D Unit of Infosys Ltd. and leading a team of 6 for Legacy Modernization initiatives under Infosys Centre for Emerging Technology Solutions (ICETS). Her skill sets are Python, Machine Learning, Front End technologies (Angular JS, HTML and JavaScript) and SQL DB. Her main research interest includes Machine learning, AI, Artificial Neural Networks. She had received her Bachelor of Technology degree from National Institute of Science and Technology (NIST), Berhampur, BPUT, Odisha in 2010. She has more than 8 years of experience is software development.