IoT Enabled Health Care System with Machine Learning Analysis

by Aayush Mallick 16bce2296

IoT Enabled Healthcare System with Machine Learning Analysis

Submitted in partial fulfilment of the requirements for the degree of

Bachelor of Technology

in

Computer Science and Engineering

by

AAYUSH MALLIK – 16BCE2296

Under the guidance of

Dr. RAMESH BABU K

School of Computer Science and Engineering

VIT, Vellore.

May, 2020

DECLARATION

I hereby declare that the thesis entitled "IoT Enabled Healthcare System with Machine Learning Analysis" submitted by me, for the award of the degree of *Bachelor of Technology in Computer Science and Engineering* to VIT is a record of bonafide work carried out by me under the supervision of Dr. Ramesh Babu K.

I further declare that the work reported in this thesis has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

Place: Vellore
Date: 28.05.2020

AAYUSH MALLIK
Signature of the candidate

CERTIFICATE

This is to certify that the thesis entitled "IoT Enabled Healthcare System with Machine Learning Analysis" submitted by **AAYUSH MALLIK (16BCE2296), School of Computer Science and Engineering,** VIT, for the award of the degree of *Bachelor of Technology in Computer science and Engineering*, is a record of bonafide work carried out by him under my supervision during the period, 01. 12. 2019 to 30.04.2020, as per the VIT code of academic and research ethics.

The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university. The thesis fulfils the requirements and regulations of the University and in my opinion meets the necessary standards for submission.

Place: Vellore

Date: 28.05.2020 **Signature of the Guide**

Internal Examiner **External Examiner**

Head of the Department
SCOPE

ACKNOWLEDGEMENTS

The idea of the project kindled when I read about the infrastructures of health care system in my country, Nepal. We definitely were a step back and I wanted to think of a project that would help propel it towards a better direction. Here, Dr. Ramesh Babu K played an integral role, as my guide and as my companion. Without him I would have lacked idea on how to approach with my project. I would like to thank him for making it possible.

I would also like to thank my parents for continuously encouraging me even though I was not able to succeed with my initial attempts. The project would not have been successful if it were not for VIT to provide me with a ground to work on. Last but not the least, I would like to express my sincere gratitude towards all my friends who were there to share their ideas and opinions to bring light to the final project.

Aayush Mallik

Executive Summary

The core concept of the project is the implementation of a healthcare system that is able to collect data from the patients and project it to the allocated health officials with a detailed analysis.

The world is going through a global pandemic right now and according to reports (Yale Studies: CNN), half of the deaths in 14 states of US comes from elderly home facilities. This has been the major factor of this report. The proposed system is designed for the weak and the elderly, the challenged and the less fortunate ones. In Nepal, the healthcare system is not managed properly. The kits that are being used in the current situation are all imported from China, and as the reports suggests 70% of them are faulty.

With such situation at hand, the proposed system uses equipment that can be easily bought and implemented as well as mass produced. The idea here is to have a system that can be implemented easily in places where it is needed the most, like the elderly home. Through data collection and analysis, a report then can be submitted to assigned health officials and also be updated regularly (real time).

CONTENTS

List of Figures

List of Abbreviations

HRM	Heart Rate Monitors
ECG	Electrocardiography
PET	Portable Electronic Thermometer
RTD	Resistance Temperature Detector
PCB	Printed Circuit Board
ICSP	In Circuit Serial Programming
OTA	Over the Air
SoC	System on Chip
GSM	Global System for Mobile
API	Application Program Interface
SMS	Short Message Service

1. INTRODUCTION

1.1. Objective

The major objective of the work is to make health system availability abundant to the rural parts majorly for the elderly and the physically challenged population. With the incorporation of computer science applied with the pre-existing healthcare system the problem of un-attendance even when the health is vulnerable would be checked upon. The general idea of the monitored health data being uploaded over cloud for the doctors to see makes it easier for them to attend the patient when the conditions prove abnormal before nick of time.

With currently corona virus on the rise, the rural elderly population would be left un-attended if they do not have anyone to look after them or are too old to visit health posts every day. The current global pandemic highlights this issue more. The number of cases is increasing and proper deployment of kits haven't been assured. The proposed system does not check for the virus in question but the concept could be used to deploy the kits making it viable for people who could not easily access it. With the proposed system, the doctors would be able to tell if the patient is sick and attend their house even when they are not physically able to. The proposed system also uses the concept of machine learning analysis. The data collected by the proposed system further uses the Machine Learning (ML) concept to give the doctor/ health post worker a detailed analysis of the collected information.

1.2. Motivation

The healthcare system of Nepal is not regulated well. The rural areas suffer the most from un-attended health checkups. For instance, the infant mortality rate is 13.2% in the rural area whereas only 6% in the urban [1].

According to a report published by *Department of Social Work* of *Tribhuvan University*, about 75.8% of the elderly population live in villages [7]. This is an alarming number considering that the population of the rural areas are not attended well by the health officials. With the pandemic in the rise, according to the reports from Yale about half of the deaths of the 14 US states come from elderly home [11]. This has been attributed to fear and/or negligence from the staffs at the elderly home. This case for Nepal would have been much worse. There is a lack of proper healthcare infrastructure and on this, the rural areas are least attended. Here, lies the motivation for this proposed idea. The elderly population are left behind mostly to look after the house whereas the offspring(s) leave for either the city or foreign soil. This leaves the elderly population vulnerable. With the lack of proper infrastructure, they are already deprived of proper healthcare and with an aging body it becomes near impossible to get the required weekly/monthly health checkup even when the health is deteriorating. This goes for the physically challenged population too. Without proper guidance they cannot be tested from their home.

Thus, the idea of the project is to implement systems that make it easier for the doctors to know about the areas of abnormality and attend the sick at their home, if the health condition passes certain threshold and with the help of ML analysis, it becomes easier to analyze the area in a whole and deploy any sources if necessary.

1.3. Background

Heart Rate Monitors (HRM) are monitoring device that are used to measure heart rates of the user (in real time). The measurement of electrical heart data can be quoted as Electrocardiography (ECG). The very early concepts utilized electrode leads attached to the user(s) chest to collect information. This idea was then broadened to be wireless and the first wireless ECG device was developed by *Polar Electro*. They utilized the model to record the heart rate for *Finnish National Cross-Country Ski* team. The idea gained quite a popularity and was further commercialized in 1983.

Likewise, a Portable Electronic Thermometer (PET) was first invented in 1954 and contained a Carboloy thermistor. They were designed to measure the temperature of the body using a thermistor or an RTD probe. They have shown to have an accuracy of +-0.2° C.

These are few of the sensors that have been implemented here in the purposed system. A smart watch is able to collect the above information with an in-built sensor but this is where the proposed system branches out. Along with the essential heart rate and the temperature measurement of the user, the system uses a Blood Pressure (BP) sensor. Companies like *Omron©* have tried to deploy the same concept into their product as well. They have devices that is able to record the blood pressure of the user(s).

These lay background for data-collection in the proposed IoT model. Now, to discuss about the concept of storing these collected data, cloud storage is deployed. It has been believed to be invented by *Joseph C. R. Licklider* in the 60's with his work on *ARPANET* to connect user(s) with the data anywhere and at any time. But it was not until late 1994 when *AT&T* provided online platform for personal and business consumption with the launch of *PersonaLink Services* that cloud storage kicked in.

The concept or the background to this proposed system is based upon the above sensors. The world is changing and trying to move all the valuable information to the cloud, so that they can be accessed anywhere and anytime. The proposed system integrates the above sensors as a back-bone and implement the cloud data collection as well.

2. PROJECT DESCRIPTION AND GOALS

According to report published by WHO (2007), the healthcare system of Nepal is concentrated in the urban areas and the rural areas are mostly neglected. The main goal of the project is to create a system that can easily be deployed to collect data from patients in both urban and the rural areas (major emphasis on the rural communities, especially the old and the physically challenged) and store it onto the online health portal. After the collection of the data, applying Machine Learning Analysis (K-Neighbors-Classifier) to get detailed analysis of the patient(s) and deploying any health worker if necessary.

The healthcare infrastructure in Nepal is average at best. People living in the rural areas are still deprived of proper daily/monthly checkup(s). The elderly population are most affected by it. It is also challenging for the physically challenged people to get their regular checkup(s). The implementation of the proposed system is to rectify this problem and establish an environment that enables the target population to have a more substantial access to the health infrastructure.

The proposed system uses an IoT based system for collecting data. Arduino Uno is the heart of the system. It is connected with Heartbeat Printed Circuit Board (PCB), Temperature Sensor and BP Sensor Module to collect the information from the user(s). The Heartbeat PCB is used to record the heart rate, the Temperature Sensor the temperature and BP Sensor Module the blood pressure of the user/patient. The collected data has to be sent over to a cloud system to be further analyzed. This is done with the help of Wi-Fi Module (ESP-01). It helps to send the data over to *ThingsSpeak* where it is further analyzed. Following diagram details the whole process:

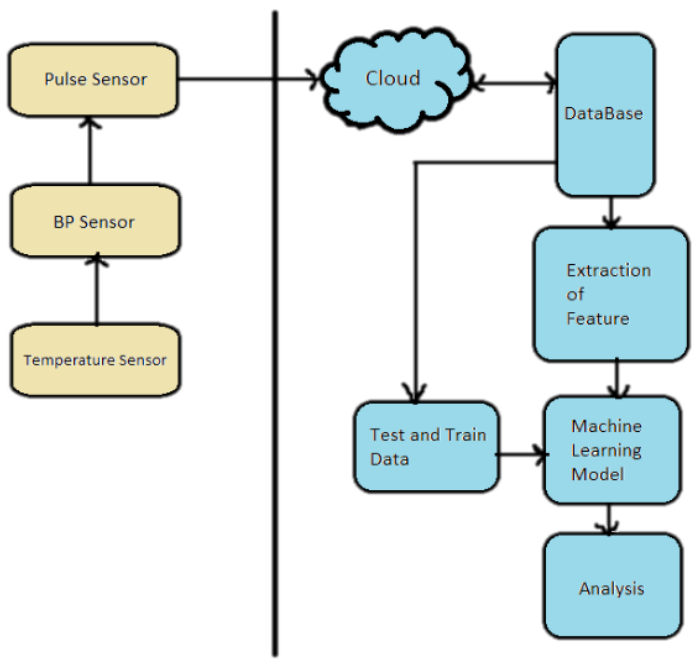

Figure 1: Diagram detailing the working process of the proposed system

The end goal here is to have a system that updates in real-time and provides health information of the user(s)/patient with minimum hassle and with the help of ML analysis, provision of detailed insight about the collected data to the health post employee(s).

3. TECHNICAL SPECIFICATION

For the technical specificity it can be broken down into two major categories. The hardware that has been used and the software that binds it all together. The proposed system is hardware extensive being an IoT device. The hardware integrated for the proposed system are:

3.1. Hardware

3.1.1. *Arduino Uno* – Arduino Uno is a microcontroller board. It consists of 14 digital input/output pins. Along with the comes a 6-analog input(s), a 16 MHz ceramic resonator, an ICSP header, a power jack, a USB connection and a reset button. Basically, it incorporates everything required to support the microcontroller. With a simple connection of it with a computer using a USB cable or with an AC-to-DC adapter it could be started and worked with.

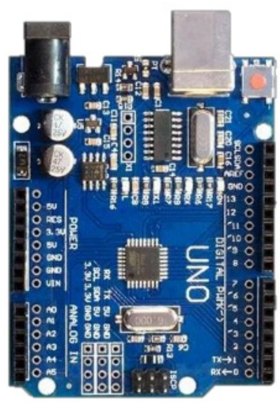

Figure 2: An Arduino Uno board

3.1.2. *Arduino Uno Wi-Fi* – Arduino Uno integrated with a Wi-Fi module results with Arduino Uno Wi-Fi. The board is schemed on ESP8266 Wi-fi Module integration. To understand it further, it is a self-contained SoC with incorporation of TCP/IP protocol stack which provides access to the Wi-Fi network. A highlighted feature of it would be the support for OTA (over-the-air) programming.

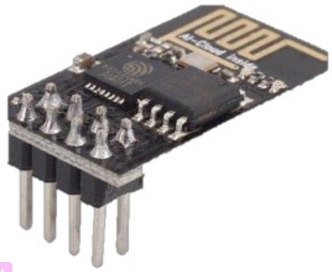

Figure 3: An Arduino Uno Wi-Fi Module

3.1.3. *Heartbeat Sensor* – Heartbeat Sensor is an electronic device that is used to measure the heart rate i.e. speed of the heartbeat. Photoplethysmography is the principle behind the working of the Heartbeat Sensor. According to this principle, the changes in the volume of blood in an organ is measured by the changes in the intensity of the light passing through that organ.

Figure 4: A Heartbeat PCB

3.1.4. *Temperature Sensor* – Temperature is the most-measured process variable in industrial automation. Most commonly, a temperature sensor is used to convert temperature value to an electrical value. Temperature Sensors are the key to read temperatures correctly and to control temperature in industrials applications.

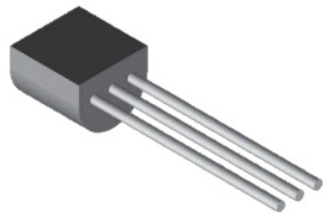

Figure 5: A Temperature Sensor

3.1.5. *BP Sensor Module* – As the name suggests, the BP sensor module is used to measure the blood pressure of the user. The data of the BP sensor module is converted into digital form and sent into to the Arduino board and then projected on through the Wi-Fi module to store it online.

Figure 6: A BP Sensor Module

3.2. Software

3.2.1. *Arduino Genuino* – The software is used to write code onto the Arduino hardware. The codes written here are compiled and displayed onto the single monitor output screen. It is used to set functions to the pins.

3.2.2. *Jupyter Notebook* – The *Anaconda Navigator*'s provision of *Jupyter Notebook* is used for applying and displaying the machine learning algorithm for the given dataset. Here, it is implemented upon the dataset taken from *Kaggel*.

4. DESIGN APPROACH AND DETAILS

4.1. Design Approach/ Materials and Methods

The proposed system would consist of the following components, pulse sensor module, blood pressure sensor module and temperature sensor module. The idea of the project is to integrate these modules to collect data from the user. All of the modules would be connected using an Arduino Uno board. It would also consist of the wi-fi module. The wi-fi module would be used to send the data onto the cloud. From here, the collected data can then be processed and analyzed using the machine learning. This would come helpful in predicting how likely a place is to be affected by a disease and thus encouraging steps to create more health posts for the betterment.

4.1.1. Functional Requirements

- Product Perspective

  The product is used for collecting data from the patient and uploading it onto the cloud for information and analysis.
- Product Features

  The product implements sensors to collect data and then uses Wi-Fi module to upload it to the cloud. The data are processed in such a way that when being written into CSV the abnormality is known.
- User Characteristics

  The user has to just give the input while the rest is assured by the proposed system.
- Assumption and Dependencies

  The assumption is that the users have a good knowledge about using the input devices and a basic knowledge of keeping the device connected at all times.

- Domain Requirements

  The proposed project would require an established net connection on both the senders end as well as the receivers end for real-time data transfer.

4.1.2. Non-Functional Requirements

- Efficiency

  The efficiency of the proposed work depends on the proper maintenance of the sensors and the strength of the Wi-Fi.
- Reliability

  The reliability of the proposed work depends on the maintenance of the sensors. It is not fragile but should be kept properly for more accurate results.
- Portability

  The system is portable with certain cautions should be taken over handling the sensors.
- Usability

  The system could serve a day-night usability without being worn out if handled properly.

The system model design details in more about the design approach of the proposed system.

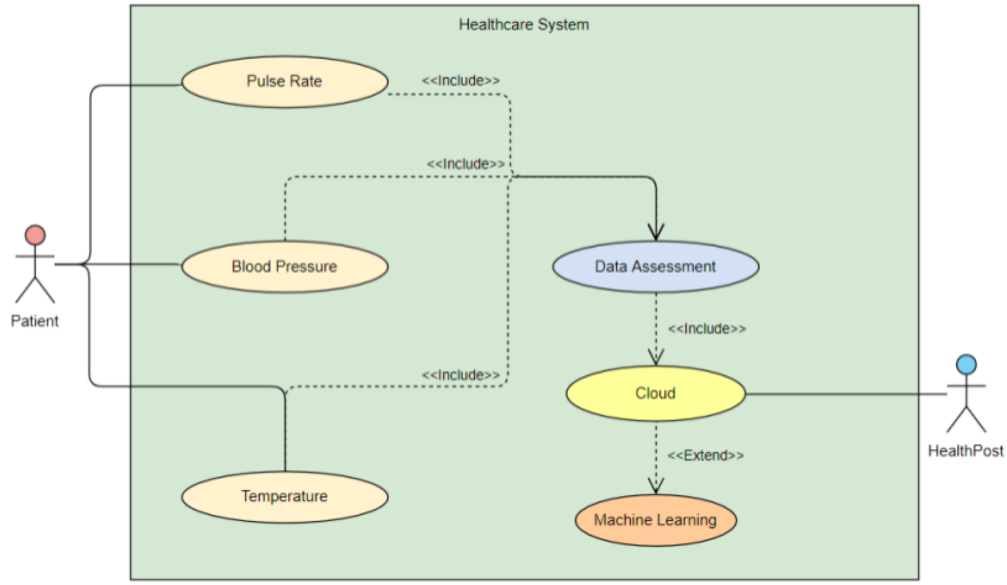

Figure 7: System model of the proposed project

The system model clearly explains and details the design approach of the proposed system. A healthcare system is established with the patient giving the data of the pulse rate, blood pressure and temperature. This data is then collected through the IoT system and sent over to a cloud storage. The data is then assessed through ML analysis and a detailed report is handed to the health post.

4.2. Codes and Standards

Codes and standards basically deal with the where what and how of the project. The codes describe the where and what aspect of the project where as the standard talks about the how portion of it. With this is mind let us look into the codes and standards of the proposed system.

- Where – In rural areas and to the people with less accessibility of the healthcare infrastructure along with the physically challenged groups.
- What – An IoT based system that is able to collect data from the user(s) and send it over to the cloud for analysis.
- How – With the help of Arduino and sensor modules for data collection and the Wi-Fi module for transmission of data. *Jupyter Notebook* than can be used for data analysis.

4.3. Constraints, Alternatives and Tradeoffs

4.3.1. Constraints

The major constraint of the proposed idea is the lack of communication infrastructure to modulate the whole process in the rural parts of the country. The whole motivation behind the project came from the lack of proper healthcare infrastructure in the rural areas but then again it is not the only thing the people there are deprived of. But a lot of initiatives are being taken to strengthen the communication bridge.

This comes as a constraint because the whole idea of the proposed system is to send collected data over to the cloud for further inspection. If the system could not access a proper internet source then it would not be able to send real-time data to the health sector operatives.

The other constraint that the project faces is with the deployment of the system. The system could be produced at large and installed as required but what about the system management? The proposed idea uses Arduino board and the wiring of the sensors are very sensitive. With any fault in the system, the health sectors would be aware of the situation through the abnormality of the receiving data but fixing and maintaining would be a task at hand.

The project suggests the implementation of the idea in a larger scale. This has not been done before and that itself is a constraint. The skepticism outweighs the potential because of the lack of test. So, to minimize this, the project has to be deployed in a smaller scale at first and then expanded upon. The resolution of the above discussed constraints and alternatives will be touched below.

4.3.2. Alternatives

The major constraint that was discussed was with the communication bridge. If the proposed system is not able to send in the data over to the cloud due to lack of proper internet service, the system would fail. But there is an alternative. Although the communication gap is being shorter with the deployment of mobile equipment by the telecommunication companies an alternative has to be sorted out. This comes with the use of GSM module.

Internet is not readily available to all the parts of rural areas but sim network signals are. The replacement of the Wi-Fi module with the GSM module would thus play a big role. Even though the internet is not readily available, the data can still be collected through SMS. The process would be hectic because the data would then have to been manually stored into the database but it is the most viable solution where there is a lack of proper internet infrastructure.

4.3.3. Trade-Offs

The tradeoff for the deployment of the alternative solution using the GSM module would have to do with the loss of capacity of bulk self-cloud data-entry. With the installation of the Wi-Fi module all the collected data were sent directly into the cloud and collected over at *ThingSpeak*'s website. But if the infrastructure lacks to hold in the internet connection, GSM module would have to be used. This would cut away the time value. If GSM module is deployed, the data would be collected over in text messages. Then these data would have to be typed in manually.

The other tradeoff that could be considered is the use of Raspberry Pi instead of Arduino board. The major difference would now be that for an infrastructure with an established internet connection, it would not require an extra module to connect to it. Raspberry Pi has that as a built-in feature but the catch is it becomes difficult to operate. Usually, Arduino is considered to be fairly easier when compared to Raspberry Pi. But if it could be compromised than it would be a good tradeoff.

5. SCHEDULE, TASK AND MILESTONE

The project started out in the early December, as soon as the previous semester ended. The initial phases of the project went into generation of ideas and analyzing real world problems that could be tackled with the implementation of computer science. Then eventually it branched out to a specific goal that the project shaped to be. Given Gantt Chart demonstrates the timeline of key events of the project.

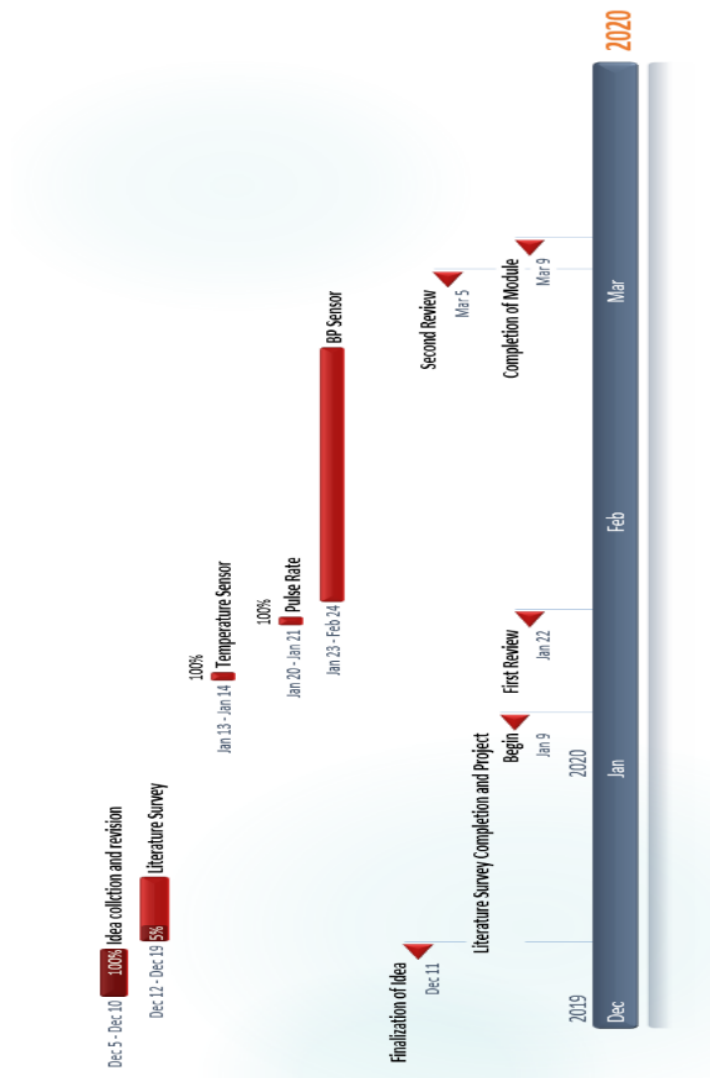

Figure 8: The key timeline of the project

The above Gantt Chart shows the timeline of the project and how it shaped. The whole process was divided into smaller tasks and worked upon. The whole thing can be broken down into the following:

- Beginning Task: The initiation

  With the very beginning of the project, the main goal was to analyze real life scenarios and come up with problems and then solution(s) that could be tackled down with the ideas of computer science. As presented in the Gantt chart this started out in the early December.
- Task of defining the problem and depth analysis of the problem

  The next task was to finalize on a particular problem and design a solution accordingly. After the problem had been defined, a depth analysis was required before the solution design. Here, literature survey was done extensively and previous research on similar topic was analyzed. Based on this, the other task was formulated.
- Task of designing the model and testing

  This was the pre-final stage of the project. A model had to be proposed that gives an answer to the problem raised. This is the physical task of the project. Over the span of more than 2 months an IoT device was designed and deployed with data collection and analysis.
- Final Task

  The final task of the project is the deployment of the concept in a larger scale. Due to misfortunates, this has not been implemented yet on a larger scale. The final task of this capstone project would remain at halt until the global pandemic begins to seize.

The milestones of the project have been highlighted in the above Gantt Chart. It highlights the dates of finalization of ideas to the formulation of the actual system along with the review dates of when it was graded and attested upon with knowledgeable panel.

6. PROJECT DEMONSTRATION

The sensors were used to collect data and then send it over to the cloud using the Wi-Fi module.

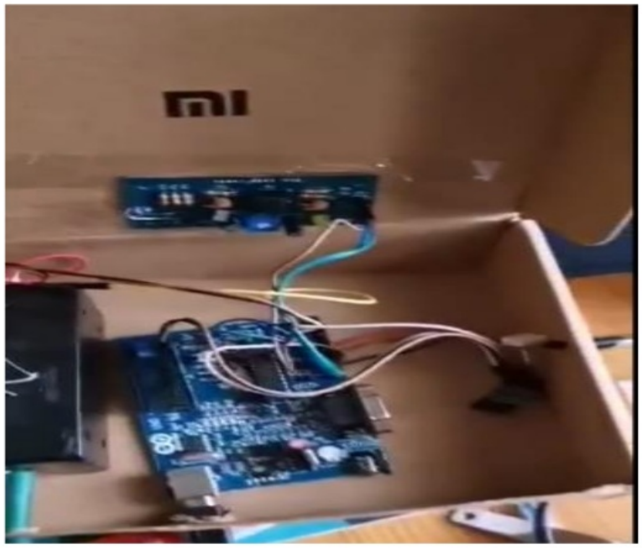

Figure 9: Physical components of the system

Thinkspeak, an IoT channel was used to host and collect the data. The API key provided for the private channel was used to link the Aruduino and the sensors to obtain the values.

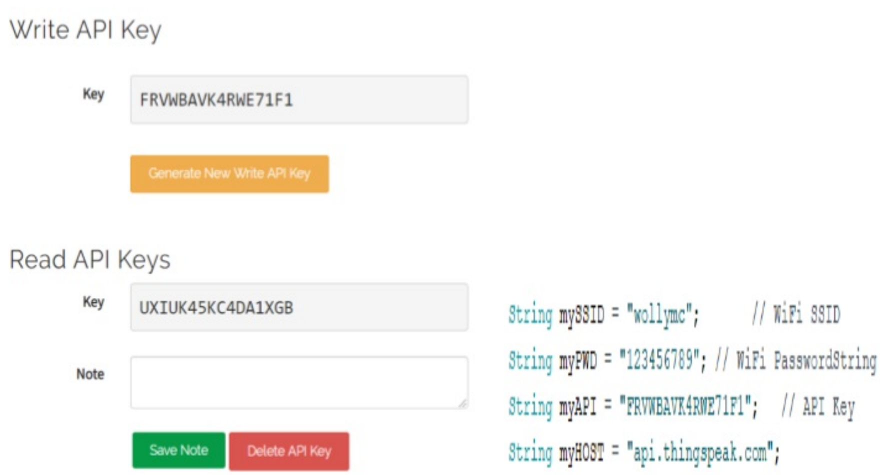

Figure 10: The write API key and the code respectively

Now the data were individually obtained for each field.

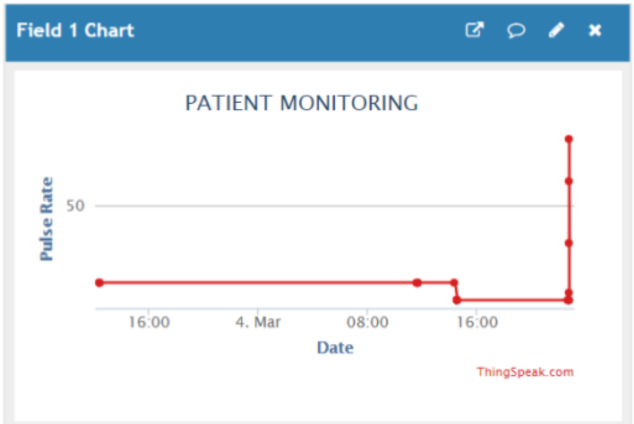

Figure 11: Field for Pulse Rate

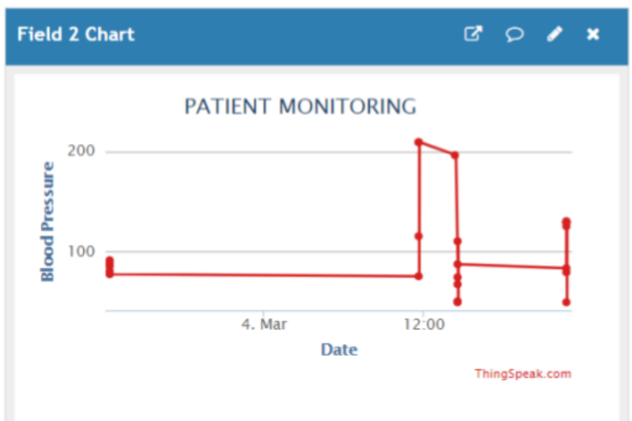

Figure 12: Field for Blood Pressure

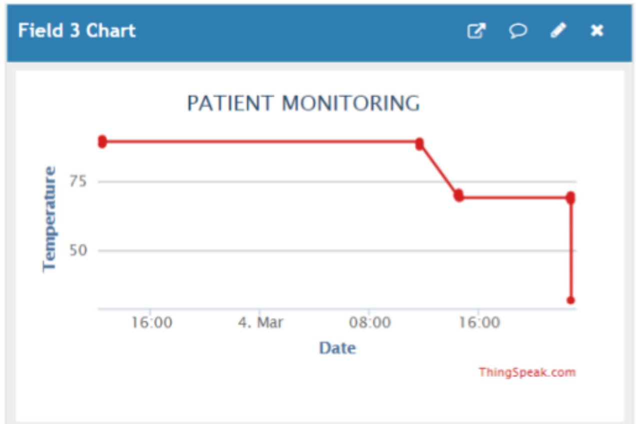

Figure 13: Field for Temperature

The following data could be then downloaded in the form of .csv, the process detailed in the section 7.3 of the document.

A1 created_at

	A	B	C	D	E	F
1	created_a	entry_id	field1	field2	field3	field4
2	2020-03-0	1	19	80	90.01	NORMAL
3	2020-03-0	2	19	91	88.25	ABNORMAL
4	2020-03-0	3	19	85	88.25	ABNORMAL
5	2020-03-0	4	19	88	90.01	ABNORMAL
6	2020-03-0	5	19	77	89.13	NORMAL
7	2020-03-0	6	19	75	89.13	NORMAL
8	2020-03-0	7	19	209	87.37	ABNORMAL
9	2020-03-0	8	19	115	88.25	ABNORMAL
10	2020-03-0	9	19	209	88.25	ABNORMAL
11	2020-03-0	10	19	196	69.79	ABNORMAL
12	2020-03-0	11	12	74	68.91	NORMAL
13	2020-03-0	12	12	49	69.79	NORMAL
14	2020-03-0	13	12	110	70.67	ABNORMAL
15	2020-03-0	14	12	50	69.79	NORMAL

Figure 14: The collected data showing normality or otherwise

Now, the collected data was not sufficient to show any depth analysis, therefore with the free dataset on *Kaggle* named *Heart Disease UCI*, K-Neighbor(s) algorithm was deployed to obtain the following analysis.

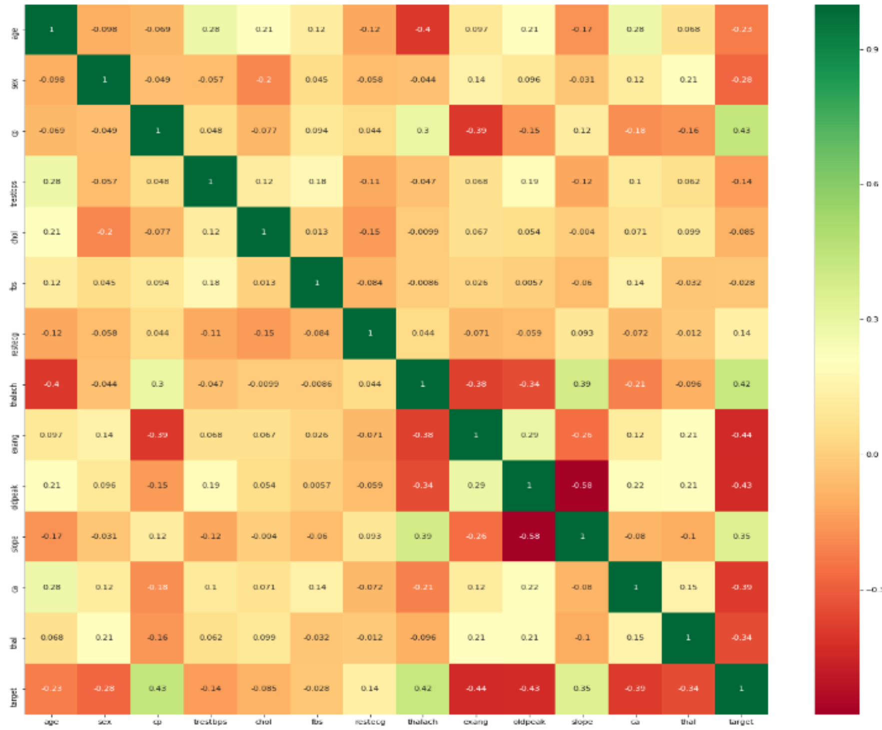

Figure 15: Chart showing the co-relation with heart disease and various symptoms

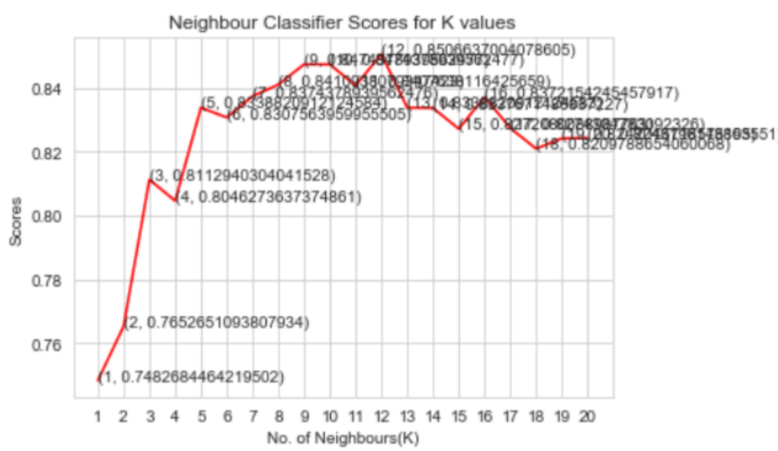

Figure 16: Graph showing the neighbor classifier scores

7. COST ANALYSIS/RESULT AND DISCUSSION

According to the reports from *Ministry of Health and Population* the budget allocation for the health sector for the fiscal year 2018/19 was NPR. 65.3bn [13]. The government is trying to push through to improve the health standards of the citizens but it isn't enough. The total cost of the proposed system was below INR 5k. Considering only few components were implemented, the cost would still be no less that INR 10k at max estimation.

The proposed idea of the project is to implement the system in a certain area. Say we try to implement in the area of density of 5 houses, each house would only require one equipment. This bring the cost estimation around INR 25k. Now, if the product is being deployed in bulk there would be a big cut in this number and through rough estimation would be around INR 18-20k. The budget allocation thus could separate a portion for the integration of this project. This way the government could ensure that everyone in that particular area is being monitored, increasing the lifestyle of the area.

The detailed analysis of the requirements for the project is attained by the points below:

7.1. Organizational Requirement

The organizational requirement stays true to its name. It could be understood as a requirement for organizing the priorities and goals of a project. The points that are to be considered for achieving this are listed as:

- Management Lead
  Here, the management lead would be the one who organizes the task and makes sure that the deployment of the units is being integrated successfully. The organization handling the project from the government's approval would be the management lead for this project.
- Structure of the Organization
  It is an important part of the organizational requirement. The structure is defined within the parameters of responsibilities and task deployment. For the project in hand, the structure would focus on production and implementation of the proposed design in the rural areas.

- Task Management

  This could basically be understood as the planning process. Task management is an important component of the organizational requirement. For this proposed system the task management would look over the structural aspect from the start till the end.

- Deployment

  The final stages of the organizational requirement involve the implementation of the proposed system. This could either be handled directly by the government or by the organization that the government assigns for the deployment of the proposed system's idea.

- Maintenance

  This is the stage of organizational requirement after the deployment of the concept. It is very necessary to monitor after the implementation. Here, the local bodies would have to take initiative to look after the deployed proposed system in the area.

7.2. Implementation Requirement

The implementation process falls under the organizational requirement but has to be separately discussed because it branches out as many important points:

- Market Assessment

  This implementation of the project requires the study of market and to see if such product already exists and/or how it could bring in any new idea to the table. For the proposed system, currently there are none deployed in the rural or the urban areas of Nepal for monitoring the health regularly of the civilians.

- Selection of Vendors

  This is an important part when it comes to the implementation of the project. As discussed, the vendor here would either be the government themselves or an organization assigned to by the government for the deployment of the project.

- Deployment

  The term deployment here is different from organizational requirement. Here, it basically means the selection of teams and assignment of roles to handle the deployment of the project.
- Research

  This process involves in the understanding of the system including the business processes before the idea goes live. Here, in the project this would be performed as a survey by the government to understand where it would be best to deploy the system.
- Testing and refinement

  This process involves the user acceptance testing as well as the refinement of the system before it is deployed completely into the market. For the project it could be done in a small scale before the idea is implemented at large.
- Finalization and support

  This is the final step of the implementation requirement. It involves the roll-out of the proposed system at large and the maintenance support provided either by an organization or the local bodies handling the installation of the project.

7.3. Results and discussions

The outputs of the project have been demonstrated in the section 6 (Project demonstration). Here, a detailed analysis on the output would be presented. So, the first part is the output from the *Thingspeak* data collection.

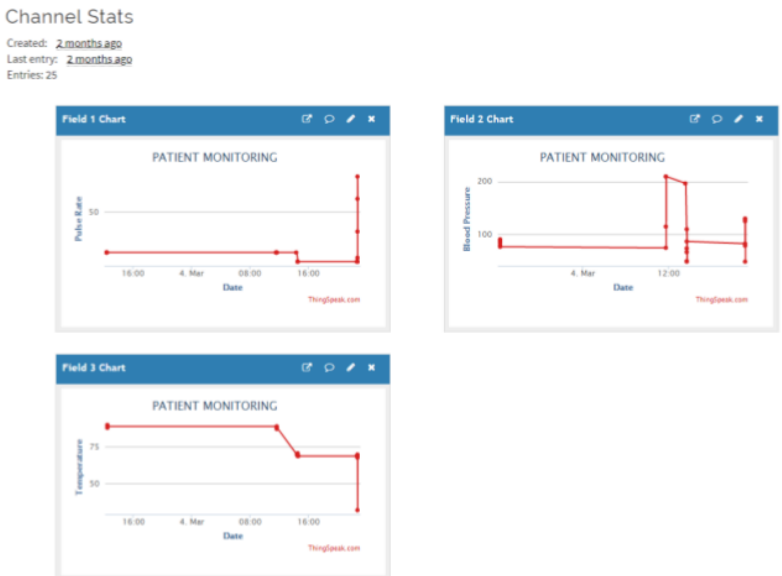

Figure 17: The output channel from the proposed model

Here, there are 3 different channels showing the output for the data collected by the proposed model. The field one shows the pulse rate and field two the blood pressure and the third the temperature. The output(s) here have a large peak and drops because the data were being collected in real time for each and every second. Any drop in peak suggests that the system was removed from the user for that particular second. The peaks having both high and low attributes to when the system was closed forcefully. The collected output than could be downloaded by exporting this data into a .csv file.

created_a	entry_id	Temperature	Pulse Rate	Blood Pressure
2020-03-0	1	19	80	90.01
2020-03-0	2	19	91	88.25
2020-03-0	3	19	85	88.25
2020-03-0	4	19	88	90.01
2020-03-0	5	19	77	89.13
2020-03-0	6	19	75	89.13
2020-03-0	7	19	209	87.37
2020-03-0	8	19	115	88.25

Figure 18: The collected data in the form of .csv document

It could be seen that the data collected here are separated into columns of different field values. The temperature column has a constant value of 19 because the temperature sensor was collecting the data of the room temperature rather than the user. If the user has used the sensor then it would have shown the value of the user's temperature. Similarly, the other fields show their respective values.

The data values were not enough to carry out a detailed analysis therefore using a heart disease dataset from *Kaggle* the report explains how it could be implemented onto the final project carried out with a larger number of samples. From the dataset, the co-relation value was found out between heart disease and different attributes. Figure 15 of the report shows just that. From this is gets clear about the attributes to be taken for analysis and the attributes to be ignored.

```
In [12]: dataset = pd.get_dummies(df, columns = ['sex','cp','fbs','restecg','exang','slope','ca','thal'])
```

```
In [13]: from sklearn.model_selection import train_test_split
         from sklearn.preprocessing import StandardScaler
         standardScaler = StandardScaler()
         columns_to_scale = ['age','trestbps','chol','thalach','oldpeak']
         dataset[columns_to_scale] = standardScaler.fit_transform(dataset[columns_to_scale])

C:\Users\omen\Anaconda3\lib\site-packages\sklearn\preprocessing\data.py:625: DataConversionWarning: Data with input dtype in
t64, float64 were all converted to float64 by StandardScaler.
  return self.partial_fit(X, y)
C:\Users\omen\Anaconda3\lib\site-packages\sklearn\base.py:462: DataConversionWarning: Data with input dtype int64, float64 w
ere all converted to float64 by StandardScaler.
  return self.fit(X, **fit_params).transform(X)
```

```
In [14]: dataset.head()
Out[14]:
```

	age	trestbps	chol	thalach	oldpeak	target	sex_0	sex_1	cp_0	cp_1	...	slope_2	ca_0	ca_1	ca_2	ca_3	ca_4	thal_0	thal_1	thal
0	0.952197	0.763956	-0.256334	0.015443	1.087338	1	0	1	0	0	...	0	1	0	0	0	0	0	1	
1	-1.915313	-0.092738	0.072199	1.633471	2.122573	1	0	1	0	0	...	0	1	0	0	0	0	0	0	
2	-1.474158	-0.092738	-0.816773	0.977514	0.310912	1	1	0	0	1	...	1	1	0	0	0	0	0	0	
3	0.180175	-0.663867	-0.198357	1.239897	-0.206705	1	0	1	0	1	...	1	1	0	0	0	0	0	0	
4	0.290464	-0.663867	2.082050	0.583939	-0.379244	1	1	0	1	0	...	1	1	0	0	0	0	0	0	

Figure 19: Data values from the dataset

The dummy columns neglect the columns that are not to be considered for the analysis. Based on this data the value of k could be determined. From Figure 16, it becomes clear that the value of k must be 12 because it has the highest peak. With this information the mean value could be found out for the dataset.

```
In [21]: knn_classifier = KNeighborsClassifier(n_neighbors = 12)
         score=cross_val_score(knn_classifier,X,y,cv=10)

In [22]: score.mean()

Out[22]: 0.8506637004078605
```

Figure 20: The mean value obtained from the score

8. SUMMARY

The devices that collect data and store it onto the cloud have been around for a while now. Smart watches collect the pulse rate and some even the temperature of the user(s) but the idea is constrained to the concept of it just being a watch. IoT sector has seen significant growth but the deployment of it in the health sector not so much. The smart watch does collect data yes, but what about a detailed analysis on it. Also, it is not affordable to everyone and is limited to collecting information of only some fields.

The idea of the following project expands upon that. The idea of a smart watch limits itself to only two major data collection components. The pulse rate and the temperature. The proposed system includes a blood pressure module too. This helps to collect the information and turn it into digital output and store it onto a cloud database. But it is not limited to this. More components could be added. Additional boards could be used to implement more components into the model.

The idea boils down to the point where it could be the most effective for the elderly population of the rural region and the physically challenged one. Currently, a global pandemic has struck the world. This leaves more room for improving upon the healthcare infrastructure as well as establish and environment for elderly and the left-out people to have their health monitored regularly.

The idea expands more than this. The data collected could then be analyzed and used to predict areas that are more suspectable to any rise of diseases with the help of ML. For instance, a correlation could be checked between certain factors and then collective data could be compared with an increasing pandemic or an epidemic to predict where the population is the most vulnerable too.

This idea could be expanded more into the future with the rise of technology. A more accurate way of data collection could be established. The current proposed system deals with an idea of a small concept but an important one that the rural elderly and the physically challenged people without any support are lacking.

References

1. D.M. Jeya Priyadharsan, K. Kabin Sanjay, S. Kathiresan, K. Kiran Karthik and K. Siva Prasath, 'Patient health monitoring using IoT with machine learning', International Research Journal of Engineering and Technology (IRJET) (Mar 2019).
2. Rohan Bhardwaj, Ankita R. Nambiar and Debojyoti Dutta, 'A Study of Machine Learning in Healthcare', IEEE 41st Annual Computer Software and Applications Conference (2017).
3. S. Sarkar and S, Misra, 'From Micro To Nano: The Evolution of Wireless Sensor-Based Healthcare', IEEE Pulse, Vol. 7, No.1, PP. 21-25 (2016).
4. Ala Al-Fuqaha, Mohsen Guizani, Mehdi Mohammadi, Mohammed Aledhari and Moussa Ayyash, 'Internet of Things: A Survey on Enabling Techonoligies, Protocols and Applications', IEEE Communications Surveys and Tutorials (Volume: 17, Issue: 4, Fourthquater) (2015).
5. Farhad Ahamed and Farnaz Farid, 'Applying Internet of Things and Machine-Learning for Personalized Healthcare: Issues and Challenges', International Conference on Machine Learning and Data Engineering (2018).
6. Shiba K Rai, Ganesh Rai, Kazuko Hirai, Ayako Abe and Yoshimi Ohno, 'The health system in Nepal – An introduction', US National Library of Medicine National Institutes of Health (2001).
7. Appechha Neupane, 'Problems encountered by ageing population in rural Nepal due to foreign migration', Tribhuvan University (2015).
8. Medical Startups. https://www.medicalstartups.org/top/ai/ (accessed on February 2020).
9. PAHS. http://www.pahs.edu.np/about/about-nepal/ (accessed on March 2017).

10. Open Source Forum. https://opensourceforu.com/2019/10/machine-learning-ml-and-iot-can-work-together-to-improve-lives/ (accessed on October 2019).
11. The Guardian. https://www.theguardian.com/us-news/2020/may/11/nursing-homes-us-data-coronavirus (accessed on May 2020)
12. Kun – Hsing Yu, Andrew L. Beam and Issac S. Kohane – Artificial Intelligence in healthcare – Nature Biomedical Engineering, 2018.
13. Budget Analysis of Ministry of Health and Population FY 2018/19, 2018.

APPENDIX A

Code for the data-collection unit

```
#include <SoftwareSerial.h>
SoftwareSerial espSerial(2, 3);
#define DEBUG true
String mySSID = "wollymc";
String myPWD = "123456789";
String myAPI = "FRVWBAVK4RWE71F1";
String myHOST = "api.thingspeak.com";
String myPORT = "80";
String myFIELD ;
float TempCel;
int BPM;
int TempPin = A0;
int pulse = A1;
int bp = A2;
int BLOOD;
float TempValue;
float Temp;
float TempFarh;
String s;
void setup()
{
  Serial.begin(9600);
  espSerial.begin(115200);

  espData("AT+RST", 1000, DEBUG);
  espData("AT+CWMODE=1", 1000, DEBUG);
  espData("AT+CWJAP=\""+ mySSID +"\",\""+ myPWD +"\"", 1000, DEBUG);
  delay(1000);

void heat()
{
```

```
 TempValue = analogRead(TempPin);
  Temp = ( TempValue/1024.0)*500;
  TempFarh = (TempCel*9)/5 + 32;
  TempCel = Temp;
 Serial.print("TEMPRATURE in Celsius = ");
 Serial.print(Temp);
 Serial.print("*C");
 Serial.print("    |    ");

 Serial.print("TEMPRATURE = ");
 Serial.print(TempFarh);
 Serial.print("*F");
 Serial.println();

 delay(1000);

}
void heart()
{
 BPM = analogRead(pulse);
 BPM = map(BPM,1000,0,0,120);
 Serial.println("BPM:");
 BPM = BPM+10;
 Serial.print(BPM);
 delay(1000);
}
void blood_pressure()
{
 BLOOD = analogRead(bp);
 BLOOD = map(BLOOD,0,1023,0,250);
 Serial.println("BP:");
 Serial.print(BLOOD);
 delay(1000);
}
```

```
void loop()
{
heat();
heart();
blood_pressure();
if(BLOOD > 180 || (BPM > 80 || BPM < 40) || TempFarh > 101 )
{
  s = "ABNORMAL";
}
else
{
  s = "NORMAL";
}
myFIELD += "&field1=";
myFIELD += String(BLOOD);
myFIELD += "&field2=";
myFIELD += String(BPM);
 myFIELD += "&field3=";
myFIELD += String(TempFarh);
myFIELD += "&field4=";
myFIELD += String(s);

  String sendData = "GET /update?api_key="+ myAPI +"&"+ myFIELD ;
  espData("AT+CIPMUX=1", 1000, DEBUG);
  espData("AT+CIPSTART=0,\"TCP\",\""+ myHOST +"\","+ myPORT, 1000,
DEBUG);
  espData("AT+CIPSEND=0," +String(sendData.length()+4),1000,DEBUG);
  espSerial.find(">");
  espSerial.println(sendData);
  Serial.print("Value to be sent: ");
  Serial.println(myFIELD);


  espData("AT+CIPCLOSE=0",1000,DEBUG);
```

```
  delay(10000);
}

String espData(String command, const int timeout, boolean debug)
{
 Serial.print("AT Command ==> ");
 Serial.print(command);
 Serial.println("     ");

 String response = "";
 espSerial.println(command);
 long int time = millis();
 while ( (time + timeout) > millis())
 {
   while (espSerial.available())
   {
     char c = espSerial.read();
     response += c;
   }
 }
 if (debug)
 {
   //Serial.print(response);
 }
 return response;
}
```

Code for data-analysis unit

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from matplotlib import rcParams
from matplotlib.cm import rainbow
%matplotlib inline
```

```
import warnings
warnings.filterwarnings('ignore')
from sklearn.neighbors import KNeighborsClassifier
df = pd.read_csv('heart.csv')
import seaborn as sns
corrmat = df.corr()
top_corr_features = corrmat.index
plt.figure(figsize=(20,20))
g=sns.heatmap(df[top_corr_features].corr(),annot=True,cmap="RdYlGn")
dataset = pd.get_dummies(df, columns =
['sex','cp','fbs','restecg','exang','slope','ca','thal'])
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
standardScaler = StandardScaler()
columns_to_scale = ['age','trestbps','chol','thalach','oldpeak']
dataset[columns_to_scale] = standardScaler.fit_transform(dataset[columns_to_scale])
dataset.head()
y = dataset['target']
X = dataset.drop(['target'], axis = 1)
from sklearn.model_selection import cross_val_score
knn_scores=[]

for k in range(1,21):
    knn_classifier = KNeighborsClassifier(n_neighbors = k)
    score=cross_val_score(knn_classifier,X,y,cv=10)
    knn_scores.append(score.mean())
plt.plot([k for k in range(1,21)], knn_scores, color = 'red')

for i in range(1,21):
    plt.text(i, knn_scores[i-1],(i, knn_scores[i-1]))
plt.xticks([i for i in range(1,21)])
plt.xlabel('No. of Neighbours(K)')
plt.ylabel('Scores')
plt.title('Neighbour Classifier Scores for K values')
```

```
knn_classifier = KNeighborsClassifier(n_neighbors = 12)
score=cross_val_score(knn_classifier,X,y,cv=10)
score.mean()
```

IoT Enabled Health Care System with Machine Learning Analysis

ORIGINALITY REPORT

14%	11%	3%	14%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to VIT University Student Paper	5%
2	Submitted to Middlesex University Student Paper	2%
3	Submitted to New Jersey Institute of Technology Student Paper	1%
4	Submitted to Loughborough University Student Paper	1%
5	www.electronicshub.org Internet Source	1%
6	www.coursehero.com Internet Source	<1%
7	www.getmscode.com Internet Source	<1%
8	ethesis.nitrkl.ac.in Internet Source	<1%
9	Submitted to Sri Lanka Institute of Information Technology	<1%

	Student Paper	
10	Submitted to University of Stirling Student Paper	<1%
11	Submitted to Australian College of Kuwait Student Paper	<1%
12	Submitted to University of Sunderland Student Paper	<1%
13	Submitted to University of Bradford Student Paper	<1%
14	Submitted to University of Wales, Bangor Student Paper	<1%
15	Submitted to Higher Education Commission Pakistan Student Paper	<1%
16	dspace.bracu.ac.bd:8080 Internet Source	<1%
17	mafiadoc.com Internet Source	<1%
18	gigadom.in Internet Source	<1%
19	Govinda K.. "chapter 8 Clinical Data Analysis Using IoT Devices", IGI Global, 2018 Publication	<1%
20	Srinivas Sethi, Ramesh K. Sahoo. "chapter 8	

Design of WSN in Real Time Application of Health Monitoring System", IGI Global, 2017 <1%
Publication

21	Submitted to Bolton Institute of Higher Education Student Paper	<1%
22	Submitted to Queen's University of Belfast Student Paper	<1%

Exclude quotes	On	Exclude matches	< 10 words
Exclude bibliography	On		

IoT Enabled Health Care System with Machine Learning Analysis

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18

PAGE 19

PAGE 20

PAGE 21

PAGE 22

PAGE 23

PAGE 24

PAGE 25

PAGE 26

PAGE 27

PAGE 28

PAGE 29

PAGE 30

PAGE 31

PAGE 32

PAGE 33

PAGE 34

PAGE 35

PAGE 36

PAGE 37

PAGE 38

PAGE 39

PAGE 40

PAGE 41