Continuous Authentication using Smartwatches

Siddarth Todi	2019B2A70991P
Sujay Patni	2019B3A70575P

Introduction

- Continuous Authentication is a method of confirming a user's identity in real time.
- Our focus is on using biometric traits or behaviors to verify the user’s identity.
- It is believed that every person has unique hemodynamics and cardiovascular system.

- Smartwatches have advanced a lot in the last few years.

- Capable of providing very accurate data on heart rate, oxygen level,

  no of step and much more.

Data Collection

- Collecting accurate biometric data from bio wearables is a very important step in our project.

- Initially focussed on finding ways to extract data from smartwatches using their own applications

- Borrowed smartwatches/fitness bands from friends and family to collect the necessary data.

Data Collection

- We had collected around 8-10 datasets from different smartwatches/fitness bands, ranging from different brands like Apple, Mi and Redmi smartbands.
- These data were collected using export options in the respective band’s app present in mobile phones linked with watches.
- **Challenges**:
  - Many brands do not support data extraction due to security concerns.(e.g. Redmi, RealMe, etc). Also can’t access sensor data directly.
  - After collecting a few datasets, we realized that the dataset varies hugely from watch to watch and brand to brand, so it is not feasible for us to analyze datasets of different types.
  - Limited number of smartwatch users on campus.
- **Possible Solution:**
  - To do sample analysis on Datasets, publicly available on internet
  - To have a smart watch, and do complete analysis using its data
  - Different brands have different sensors and hence different parameters are collected from these watches. Also the format of data collection is also different for different watches.

Data Collection

1) Apple Watch
   Phone: Iphone
   App: Health
   Steps: Iphone -> Health App -> Click on Profile image on top left -> Export All Health Data -> Export -> Share via various options
   The data collected is a zip file and within it .xml files are present. We then need to convert xml to csv in Excel.

2) MI Watch
   App: MI Fit
   Steps: Phone -> MI Fit App -> Profiles -> settings -> about -> exercising user rights -> export
   The user needs to select the parameters and date, from which he/she wants the data to be exported.
   The data collected in a zip file and a password is required to open the zip file.

3) Redmi Smart Band Pro
   App: MI Fit
   Steps:
   The user needs to select the parameters and date, from which he/she wants the data to be exported.
   The data collected in a zip file and a password is required to open the zip file.

Data Collection

- Next, we researched on various sensors present in different smartwatches.

- Major sensors present in bio wearables include:
    - Accelerometer: Tracks movement of your body
    - Gyroscope: Detects motion and gestures
    - Altimeter: Detects change in height and altitude
    - Temperature Sensors: Measures body temperature as well as surrounding temperature
    - Optical Heart Rate Sensor: Measures pulse rate
    - Oximetric Sensor: SpO2 and Oxygen level
    - ECG Sensor: Heart's Rhythm and electrical activity
    - And so on

Data Collection

- Selected 3-4 smartwatches that provide the maximum number of sensor capabilities and are economically feasible to work on.

sno	brand	model	spo2	24x7 body temperature	blood_pressure	heart_rate	exercise_modes	auto_sleep_tracking
1	goqii	personal care with smart vital plus	y	y	y	y	18	y
2	goqii	personal care with smart vital	y	y	y	y	18	y
3	goqii	personal care with vital 4	y	y	y	y	17	y
4	Fitbit	Charge 5	y	y	n	y		y
5	Amazfit	Bip 3	y	n	n	y	60	y
6	Espruino	Bangle.js 2	n	y	n	y		n
7	Denver	BFH-153	n	n	y	y		y
8	Denver	BFH-252	y	n	y	y		y
9	Denver	164 BlackMK2	y	y	n	y		y
10	Pebble	zen-pro	y	n	y	y		n
11	Enhance Colmi	colmi P8 Plus	y	n	y	y	8	y
12	Ambrane	Fitshot Loop	y	n	y	y		y
13	Dr Trust USA	Healthpal 1	y	y	y	y		y
14	Fire-boltt	Mercury	y	y	n	y		y
15	Hammer	Pulse Oximeter	y	y	y	y		y
16	Fire-boltt	Talk	y	n	y	y		y
17	PineTime	Open Source, hackable	n	n	n	y		n

Data Collection

- Watches selected: Goqii Smart Vital, Hammer Pulse Oximeter and OnePlus SmartBand

- **Goqii Smart Vital**: Optical Heart Sensor, SpO2, Body Temperature and Blood Pressure along with Movement Trackers
  - Challenges: App doesn't have the necessary capabilities, No developer mode
- **Hammer Pulse Oximeter**: Optical Heart Sensor, SpO2, Body Temperature and Blood Pressure along with Movement Trackers
  - Challenges: No developer mode, Doesn't store fitness data frequently
- OnePlus SmartBand: Optical Heart Sensor, SpO2 and Blood Pressure(different from others as its a fitness band instead of smartwatch)

Data Collection

- Shifted focus on heart rate specifically, to focus on univariate models.

- Heart Rate: the number of times the heart beats within a certain time period, usually a minute.
- Collected heart rate data from smartwatches

- Challenges:
  - As the data is collected on per minute basis, only 1 datapoint per minute is available on all the smartwatches.
  - Thus the data is not useful for continuous authentication.

Data Collection

- Photoplethysmography (PPG) is a non-invasive method for optical measurement of changes in tissue blood volume.
- The basic setup consists of a light source irradiating the tissue under examination, and a detector registering changes in light intensity due to light-tissue interaction.

- Most smartwatches use PPG method to find Heart Rate of the user
- Thus, we lose a lot of features when using heart rate instead of PPG.
- Furthermore, PPG data is as frequent as 10 data points in a single second, which can provide enough
- Thus, shifted focus to PPG data instead of Heart Rate.

Data Collection

- Extracting raw PPG data from smartwatches is a big challenge
- None of the watches provide direct access to their raw data

- Worked on creating an API for Google Fit to extract PPG data from the smartwatches.
- However, none of the watches provide access to developer options to install custom applications on them. The option is only available in high end android watches.
- Looked into creating an application on android phone and then using Google Fit to access the smartwatches for data

- Couldn’t find any concrete solution for the same.

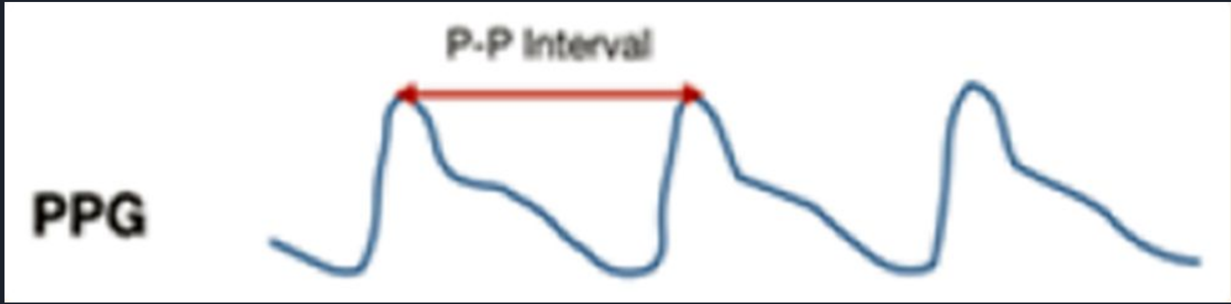
Data Collection

- Extracting raw ppg data from Smartwatches a difficult task
- Initially: aim to extract data directly from PPG sensors

- Future Work: Finding the right ppg sensors for data extraction and working with it
- SmartCare wrist-worn pulse oximeter
- Heart Pulse Monitor: Fossil Gen 5, Huawei watch 2, Galaxy watch 4 (Samsung reverted back to Wear os)
- MAXREFDES100
- HRM2511e
- MAX30100 Pulse Oximeter Heart Rate Sensor Module
- MAX30101
- Empatica E4 wristband sensor
- National Instruments device (NI cDAQ-9172)
- Maxim Integrated MAXREFDES100 device.
- SOMNOtouch NIBP
- NJL5310R, NJR Corporation, Japa

PPG Research

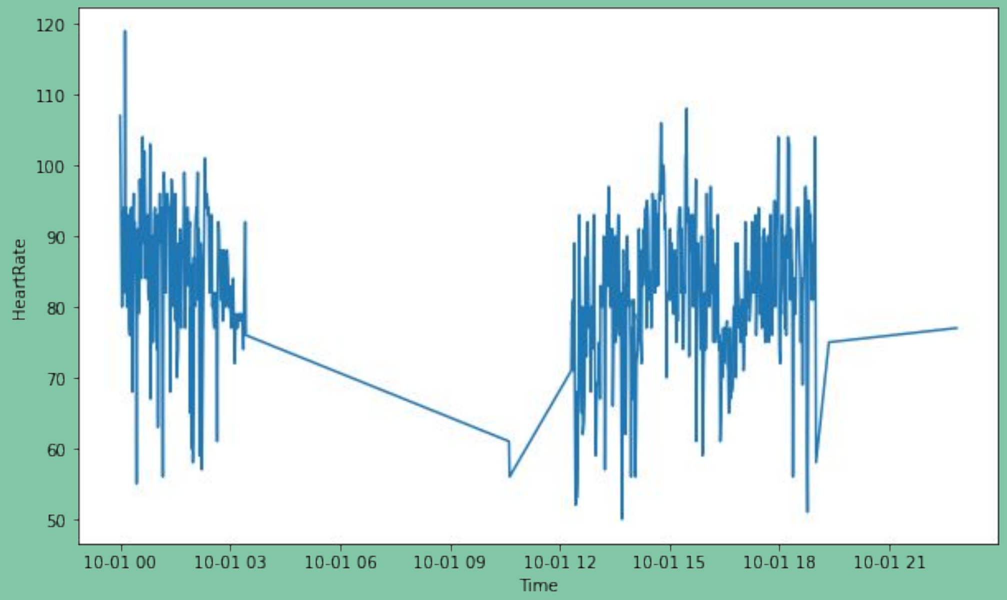
- Photoplethysmography (PPG) is a simple and low-cost optical technique that can be used to detect blood volume changes in the microvascular bed of tissue.
- The PPG waveform comprises a pulsatile ('AC') physiological waveform attributed to cardiac synchronous changes in the blood volume with each heartbeat, and is superimposed on a slowly varying ('DC') baseline with various lower frequency components attributed to respiration, sympathetic nervous system activity and thermoregulation.
- PPG sensors are extremely sensitive to motion, particularly in a wearable device, and have significant challenges measuring biometrics accurately during daily activities and exercise.
- PPG signal’s second derivative wave contains important health-related information.



Experimentation: Heart Rate

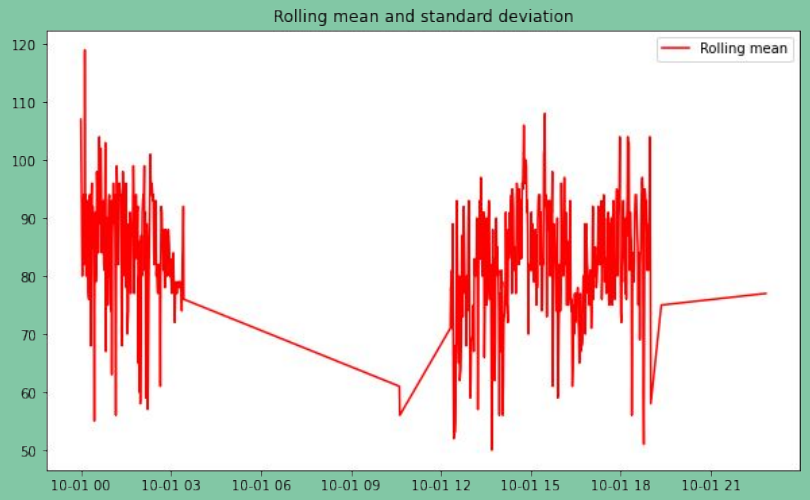
Using one of the data which we had collected, we had done time series analysis on heart rate data. (Snapshot of a part of data attached below)

1	time	heartRate
2	0:00	107
3	0:01	97
4	0:02	84
5	0:03	80
6	0:04	90
7	0:05	93
8	0:06	94
9	0:07	82
10	0:08	119
11	0:09	91
12	0:10	94
13	0:11	80
14	0:12	92
15	0:13	93
16	0:14	84
17	0:15	77
18	0:16	76

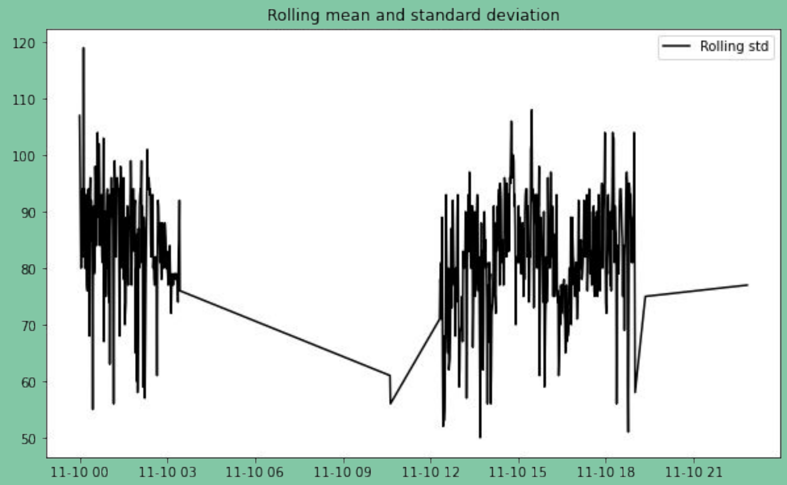


Experimentation-Statistics

Rolling Mean(Window size=60)



Standard Deviation(Window size=60)



Challenges: Scattered discontinuous data and also step wise data(1 point/minute)

Experimentation-Statistics

Dickey-Fuller test was also conducted on this data (tests null hypothesis in univariate stationary data)

```
Result of Dickey-Fuller Test
Test Statistic                  -4.723755
p-value                          0.000076
#Lags Used                      10.000000
Number of observations Used    590.000000
Critical Value (1%)             -3.441482
Critical Value (5%)             -2.866451
Critical Value (10%)            -2.569386
dtype: float64
```

Experimentation

We had taken 2 datasets of 2 hour interval each, of two different days of the same person.

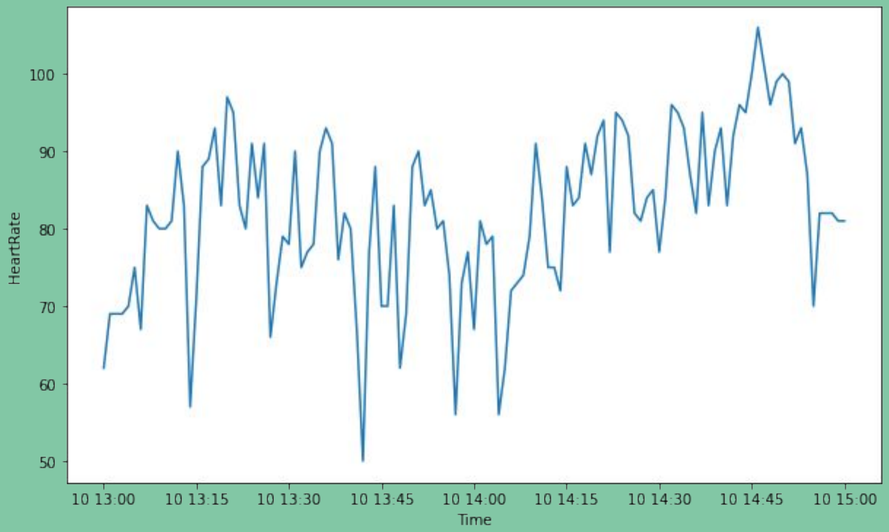
1	time	heartRate
2	13:00	62
3	13:01	69
4	13:02	69
5	13:03	69
6	13:04	70
7	13:05	75
8	13:06	67
9	13:07	83
10	13:08	81
11	13:09	80
12	13:10	80
13	13:11	81
14	13:12	90
15	13:13	83
16	13:14	57
17	13:15	71
18	13:16	88

1	time	heartRate
2	13:00	77
3	13:01	90
4	13:02	82
5	13:03	91
6	13:04	81
7	13:05	77
8	13:06	76
9	13:07	92
10	13:08	93
11	13:09	80
12	13:10	90
13	13:11	83
14	13:12	102
15	13:13	102
16	13:14	87
17	13:15	78
18	13:16	90

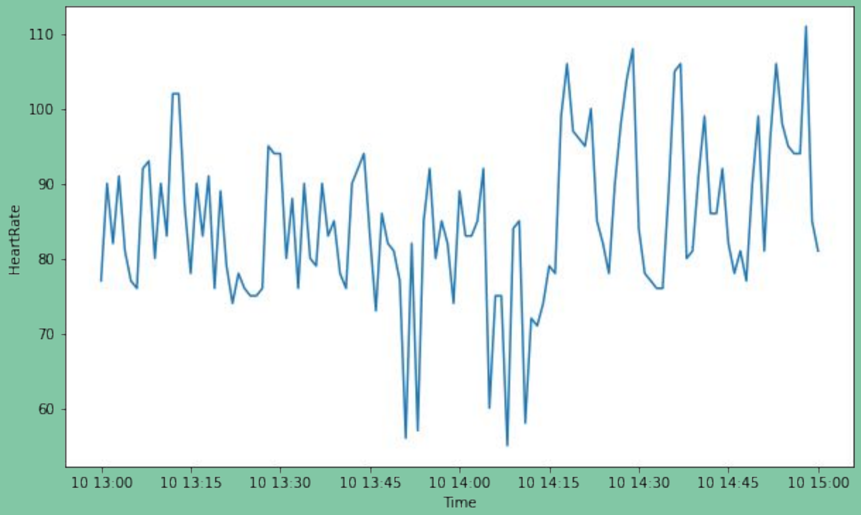
Experimentation

Plot of the 2 datasets taken of 2 hours on different days

Data 1

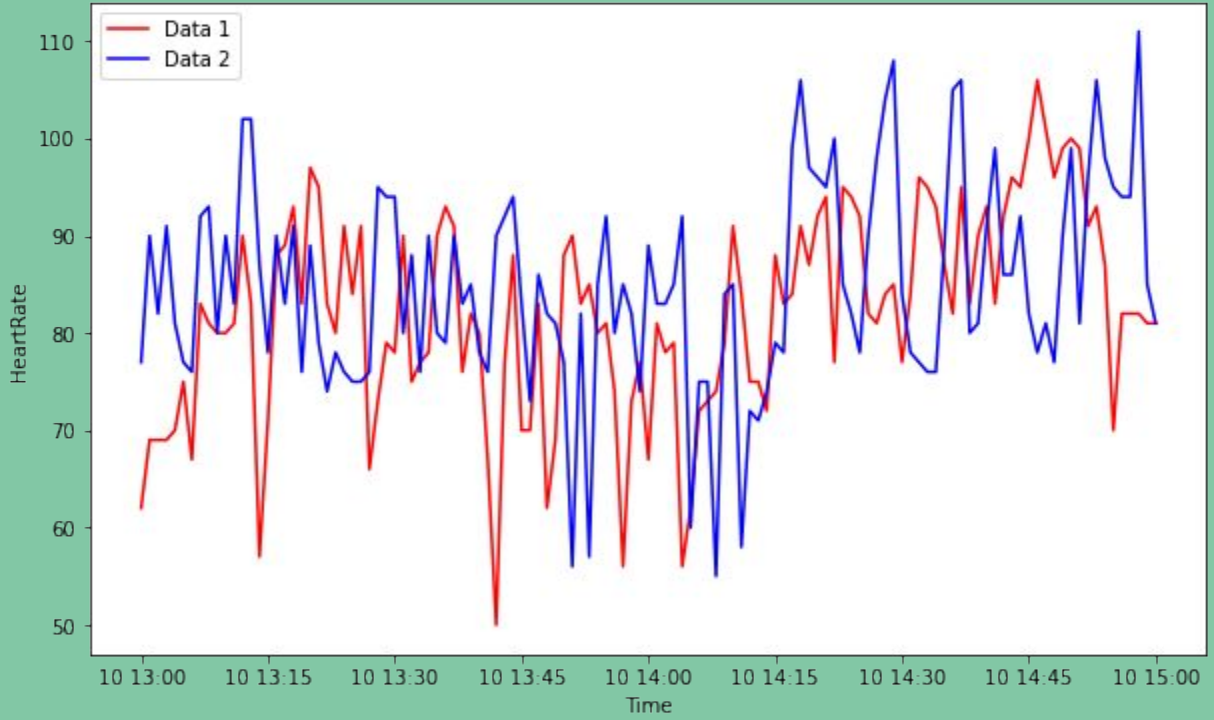


Data 2



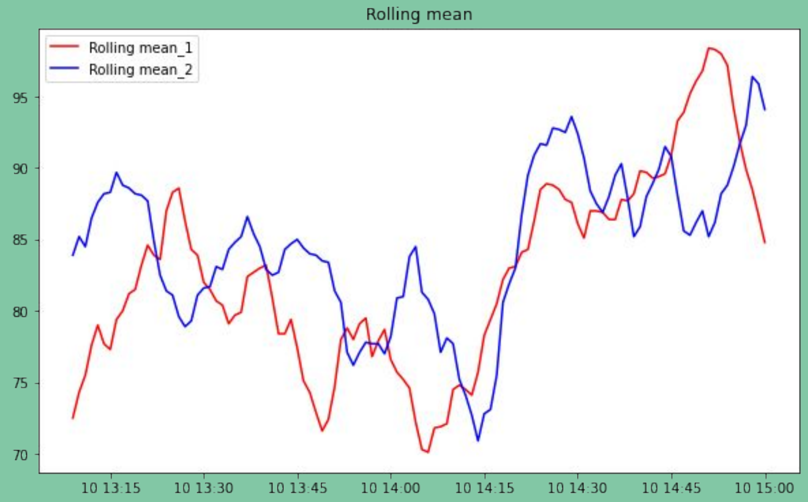
Experimentation

Comparing the plots of both the datasets

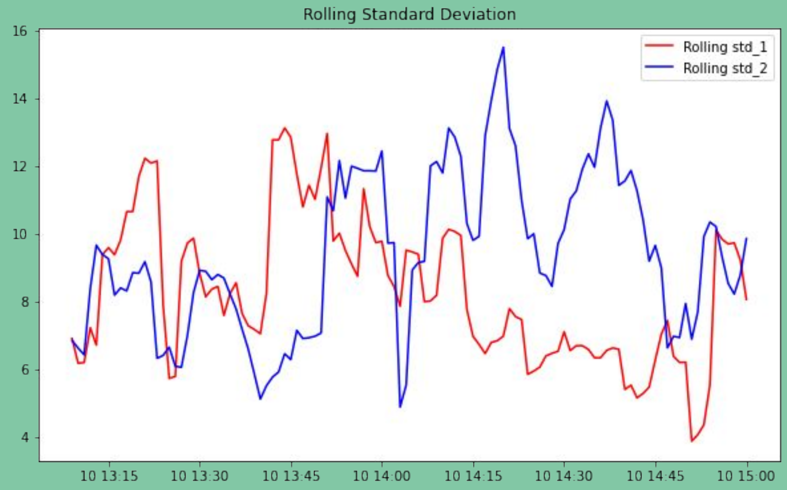


Experimentation-Statistics

Rolling Mean(Window size=10)



Standard Deviation(Window size=10)



SARIMA Model

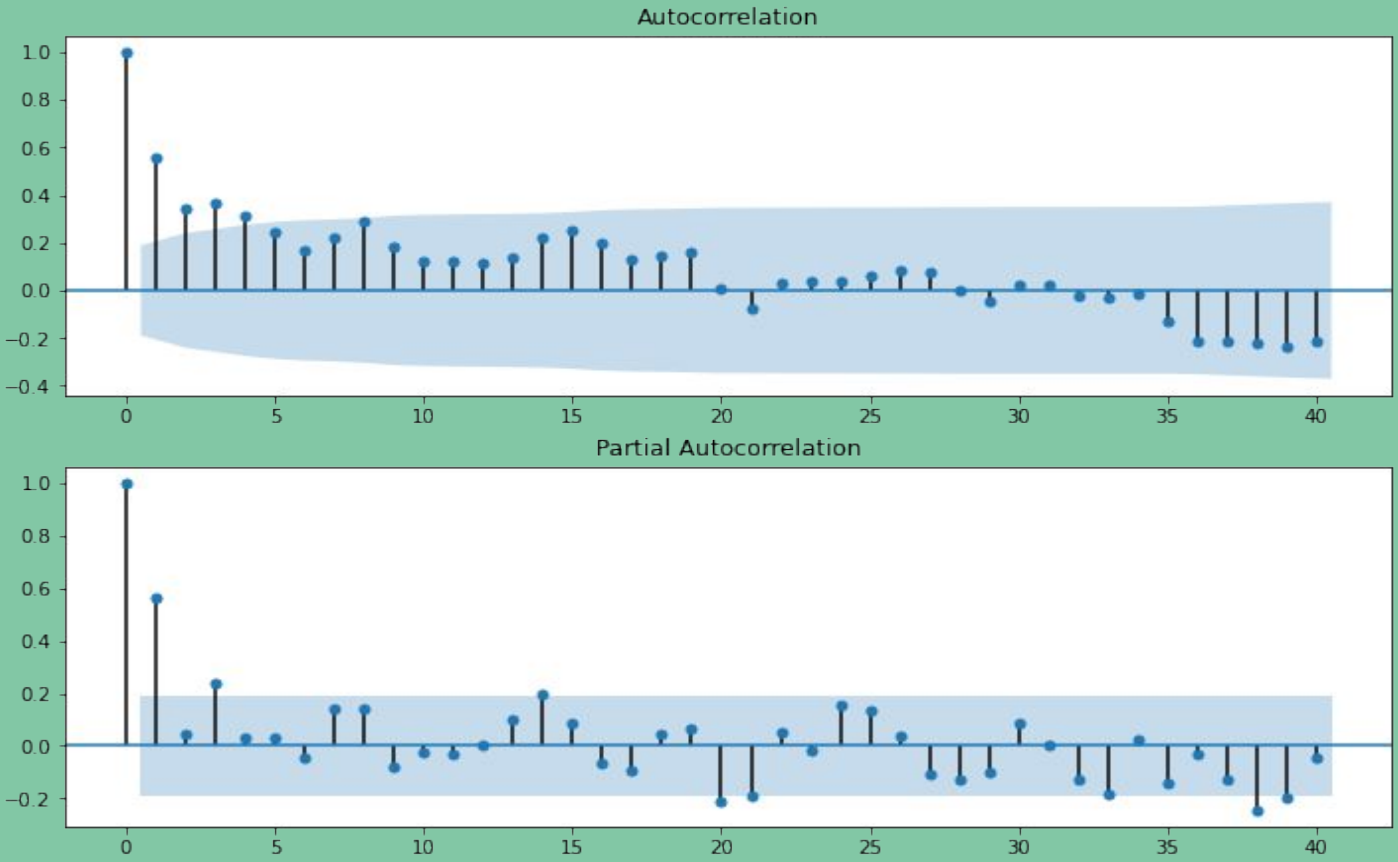
It is a time-series forecasting model implemented on univariate stationary data.

We had implemented this model, on the 2 hour heart rate data(data-1, as shown previously)

```
Test parameters : -3.587722330343093

p-value : 0.006000135418084747

#Lags in the database : 2

Data points used : 118

Dataset is stationary
```

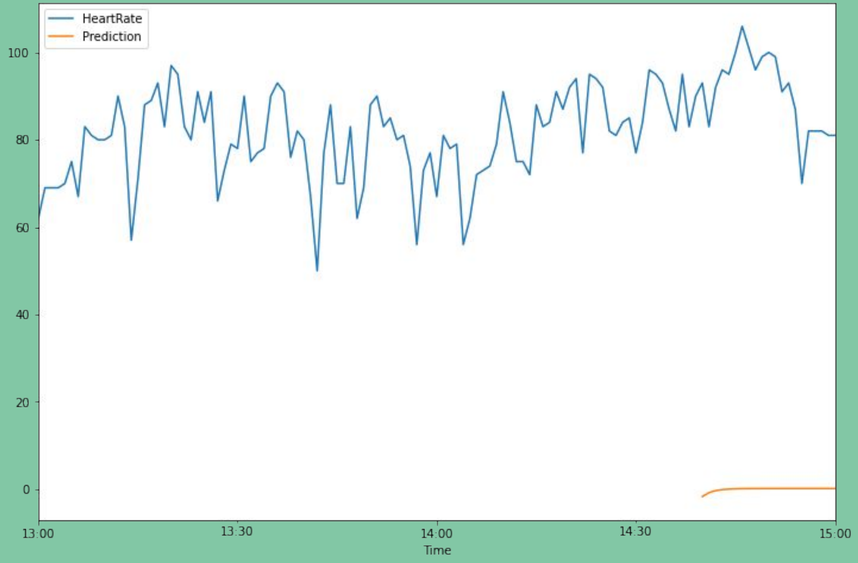
As the p-value is less than 0.05, the data is stationary, otherwise we had to do normalization.

SARIMA Model

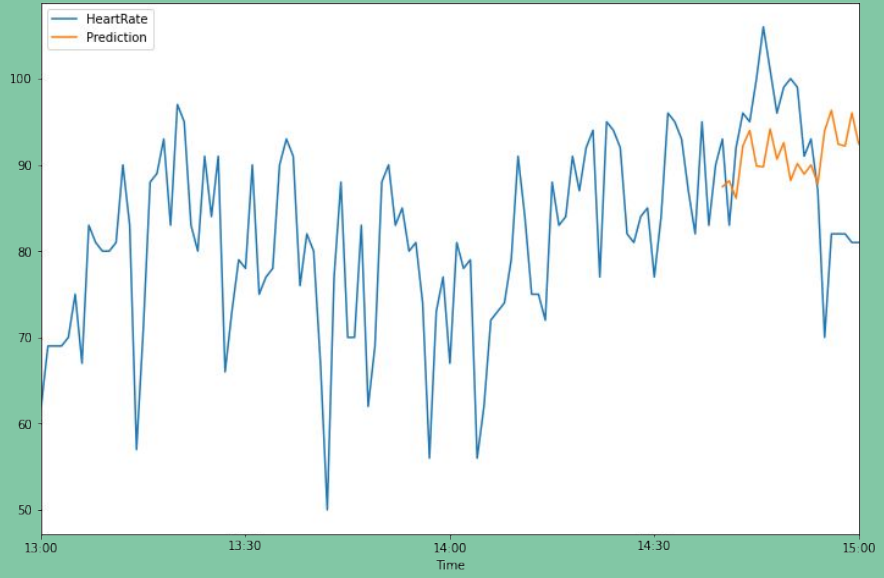


Predictions

ARIMA Prediction



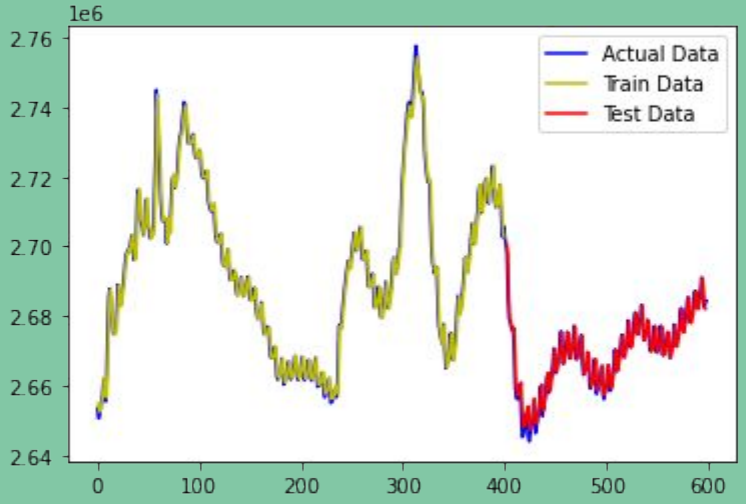
SARIMA Prediction(Slot size=12)



Long Short Term Memory(LSTM)
Type-1: 1 minute

Total number of points: 599
Train points: 401
Test points: 198
Look back: 1
`Train Score: 3683.60 RMSE`
`Test Score: 3419.09 RMSE`

	A	B
1	ppg	timer
2	2653440	10:48:00
3	2650560	10:48:00
4	2650767	10:48:00
5	2653281	10:48:00
6	2655284	10:48:00
7	2657600	10:48:01
8	2661857	10:48:01
9	2658077	10:48:01
10	2655382	10:48:01
11	2662347	10:48:01
12	2675967	10:48:01
13	2685597	10:48:01
14	2687903	10:48:01
15	2685354	10:48:01
16	2679560	10:48:01

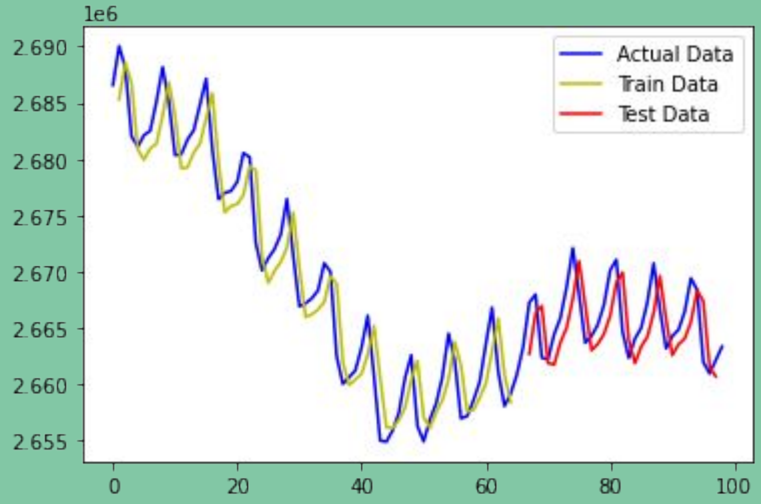


Long Short Term Memory(LSTM)
Type-2: 10 seconds

Total number of points: 99
Train points: 66
Test points: 33
Look back: 1

```
Train Score: 3015.30 RMSE
Test Score: 2990.91 RMSE
```

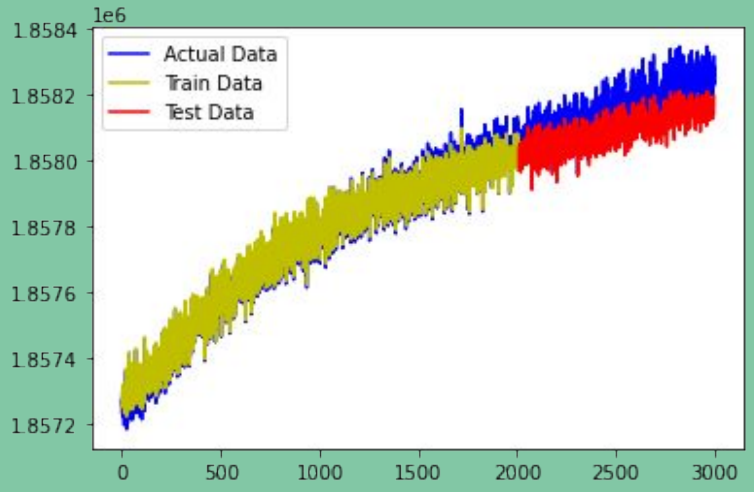
35	2668344	10:49:03
36	2670754	10:49:03
37	2670008	10:49:03
38	2662499	10:49:04
39	2660045	10:49:04
40	2660637	10:49:04
41	2661234	10:49:04
42	2663271	10:49:04
43	2666107	10:49:04
44	2660809	10:49:04
45	2655025	10:49:04
46	2654917	10:49:04
47	2655926	10:49:04
48	2657462	10:49:05
49	2660578	10:49:05
50	2662600	10:49:05
51	2656310	10:49:05



Long Short Term Memory(LSTM)
Type-3: 5 minutes

Total number of points: 2999
Train points: 2009
Test points: 990
Look back: 1
`Train Score: 65.37 RMSE`
`Test Score: 69.15 RMSE`

1979	1857923	10:54:18
1980	1857953	10:54:18
1981	1857995	10:54:18
1982	1858053	10:54:18
1983	1858058	10:54:18
1984	1858041	10:54:18
1985	1858015	10:54:18
1986	1858053	10:54:18
1987	1858126	10:54:18
1988	1858068	10:54:19
1989	1858004	10:54:19
1990	1858112	10:54:19
1991	1858004	10:54:19
1992	1858123	10:54:19
1993	1858076	10:54:19
1994	1858052	10:54:19
1995	1858017	10:54:19
1996	1858047	10:54:19
1997	1858041	10:54:19
1998	1858036	10:54:20
1999	1857987	10:54:20
2000	1858029	10:54:20



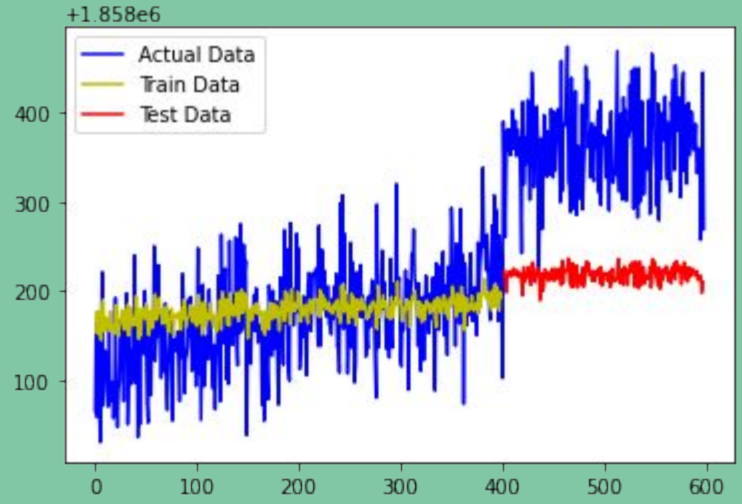
Long Short Term Memory(LSTM)

Type-4: 40(train)-20(random) seconds (Look back: 1)

Total number of points: 599
Train points: 401
Test points: 198
Look back: 1

```
Train Score: 54.89 RMSE
Test Score: 154.33 RMSE
```

395	1858211	10:55:39
396	1858291	10:55:39
397	1858224	10:55:39
398	1858166	10:55:39
399	1858262	10:55:39
400	1858235	10:55:39
401	1858243	10:55:39
402	1858102	10:55:40
403	1858390	10:56:40
404	1858325	10:56:40
405	1858260	10:56:40
406	1858353	10:56:40
407	1858381	10:56:40
408	1858357	10:56:41
409	1858381	10:56:41



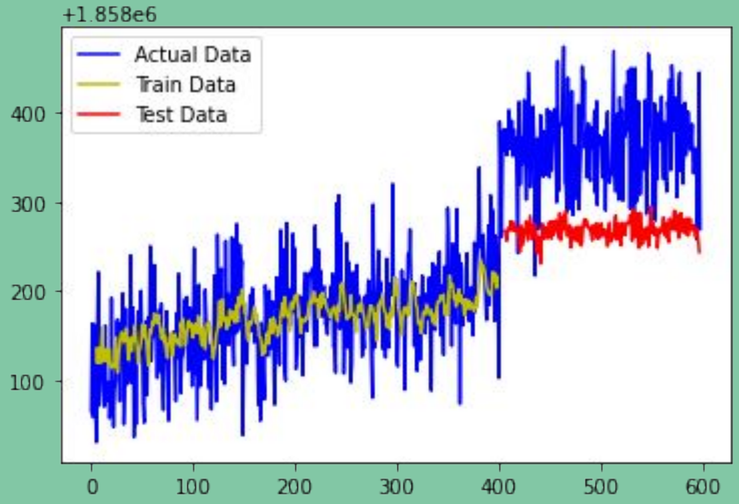
Long Short Term Memory(LSTM)

Type-4: 40(train)-20(random) seconds (Look back: 5)

Total number of points: 599
Train points: 401
Test points: 198
Look back: 5

```
Train Score: 49.75 RMSE
Test Score: 109.34 RMSE
```

395	1858211	10:55:39
396	1858291	10:55:39
397	1858224	10:55:39
398	1858166	10:55:39
399	1858262	10:55:39
400	1858235	10:55:39
401	1858243	10:55:39
402	1858102	10:55:40
403	1858390	10:56:40
404	1858325	10:56:40
405	1858260	10:56:40
406	1858353	10:56:40
407	1858381	10:56:40
408	1858357	10:56:41
409	1858381	10:56:41



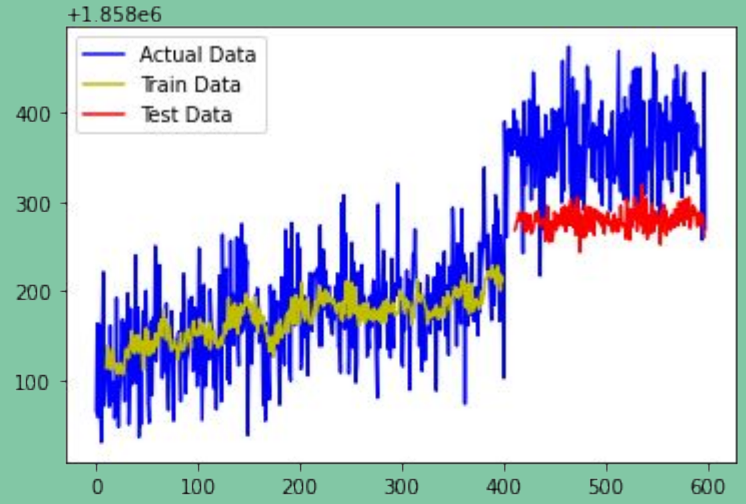
Long Short Term Memory(LSTM)

Type-4: 40(train)-20(random) seconds (Look back: 10)

Total number of points: 599
Train points: 401
Test points: 198
Look back: 10

`Train Score: 48.35 RMSE`
`Test Score: 98.65 RMSE`

395	1858211	10:55:39
396	1858291	10:55:39
397	1858224	10:55:39
398	1858166	10:55:39
399	1858262	10:55:39
400	1858235	10:55:39
401	1858243	10:55:39
402	1858102	10:55:40
403	1858390	10:56:40
404	1858325	10:56:40
405	1858260	10:56:40
406	1858353	10:56:40
407	1858381	10:56:40
408	1858357	10:56:41
409	1858381	10:56:41



Next Step

- Divide the work into 3 parts:
  - PPG Data Collection
  - Time Series Prediction
  - Matching Time Series Models

- Data Collection: Need assistance to look for PPG extraction methods and
- Time Series Prediction: Using LSTM and Prophet Model
- Matching Time Series: Statistical Approach (Gaussian Mixture, Mixture of Models ....)

THANK YOU