Continuous Authentication using Smartwatches

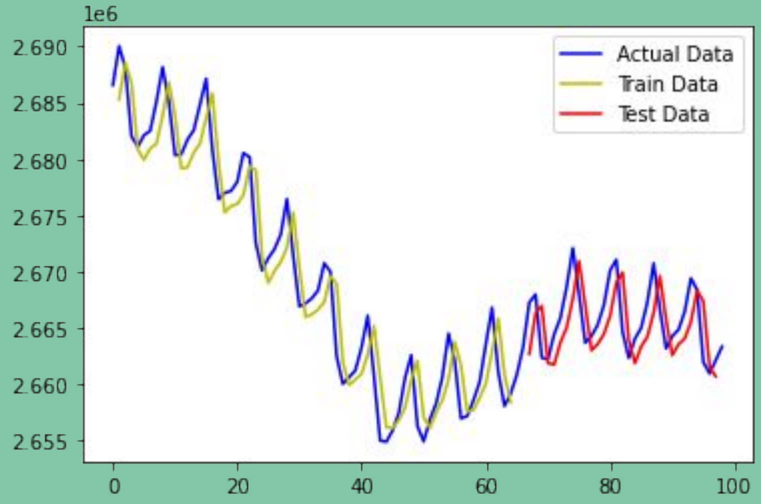
Siddarth Todi 2019B2A70991P
Sujay Patni 2019B3A70575P

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
```

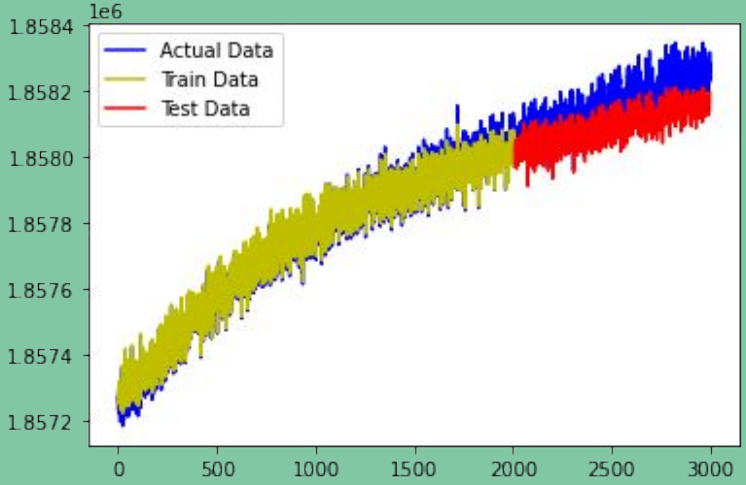
35	2668344	10:49:03
36	2670754	10:49:03
37	2670008	10:49:03
38	2662499	10:49:04
39	2660045	10:49:04
40	2660637	10:49:04
41	2661234	10:49:04
42	2663271	10:49:04
43	2666107	10:49:04
44	2660809	10:49:04
45	2655025	10:49:04
46	2654917	10:49:04
47	2655926	10:49:04
48	2657462	10:49:05
49	2660578	10:49:05
50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



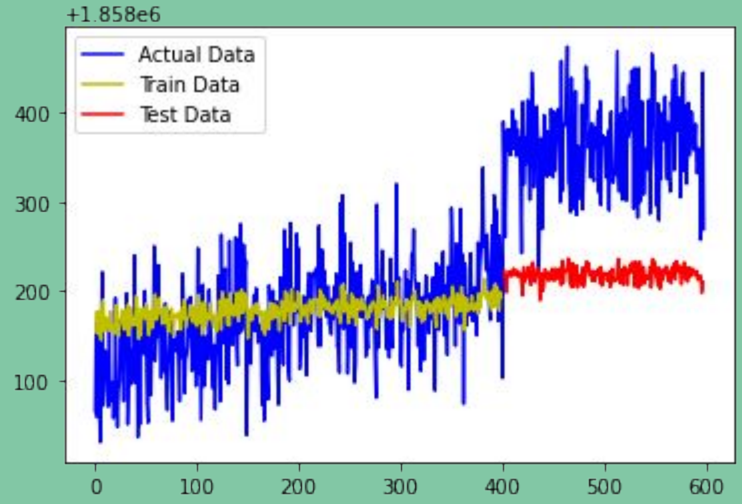
Long Short Term Memory(LSTM)

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



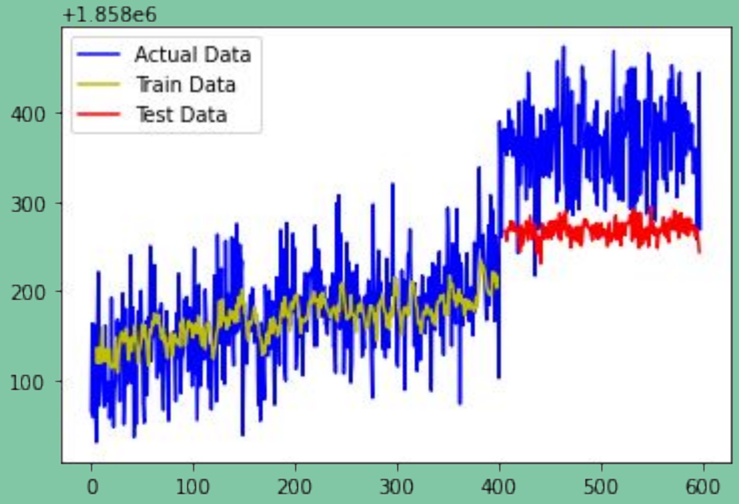
Long Short Term Memory(LSTM)

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



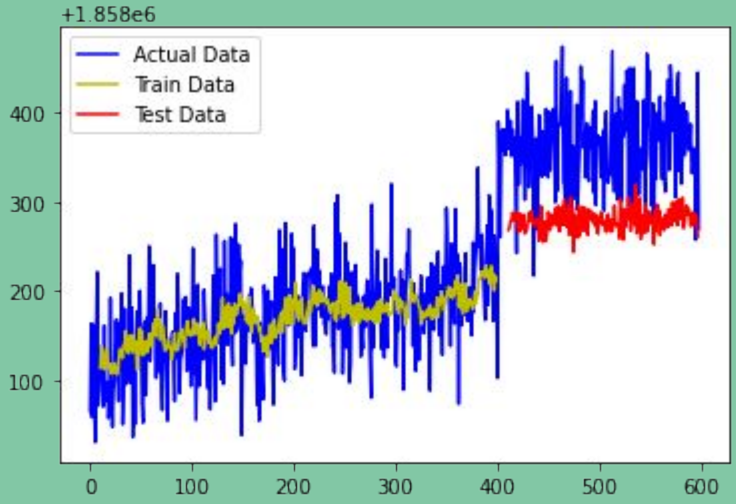
Long Short Term Memory(LSTM)

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
```

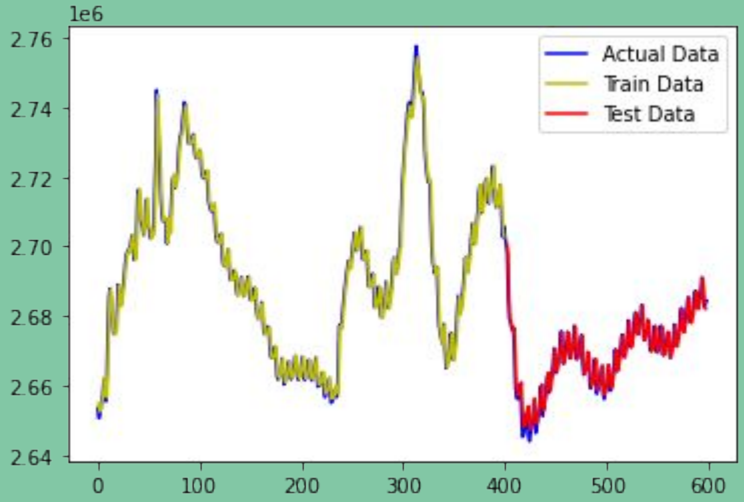
395	1858211	10:55:39
396	1858291	10:55:39
397	1858224	10:55:39
398	1858166	10:55:39
399	1858262	10:55:39
400	1858235	10:55:39
401	1858243	10:55:39
402	1858102	10:55:40
403	1858390	10:56:40
404	1858325	10:56:40
405	1858260	10:56:40
406	1858353	10:56:40
407	1858381	10:56:40
408	1858357	10:56:41
409	1858381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



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



Problem: Testing data uses real values to predict next value

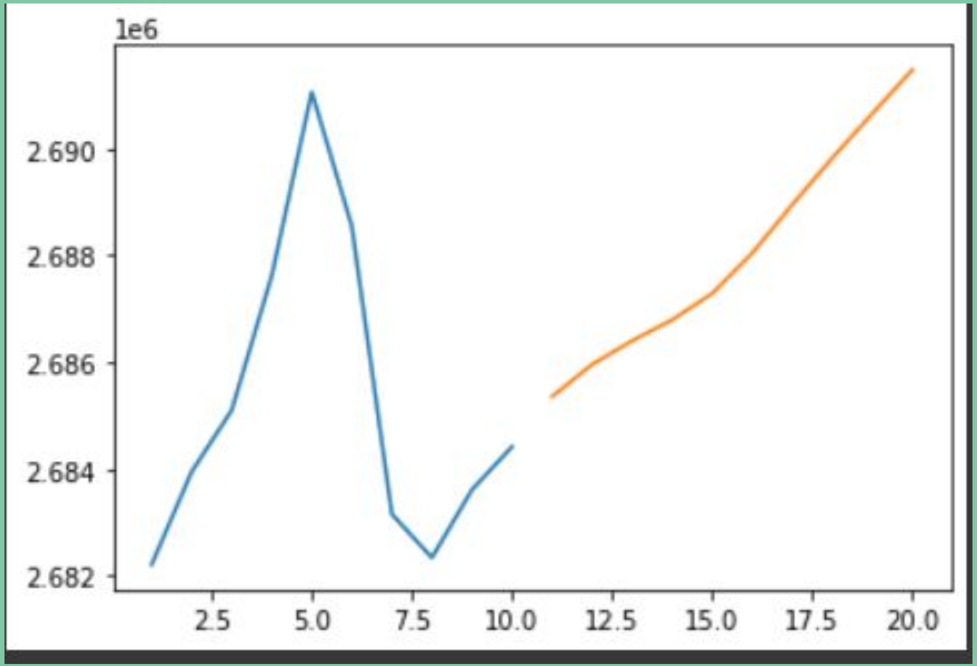
Long Short Term Memory(LSTM)

Type-1: 1 minute

Worked on code that uses predicted values instead of real values to predict the next data

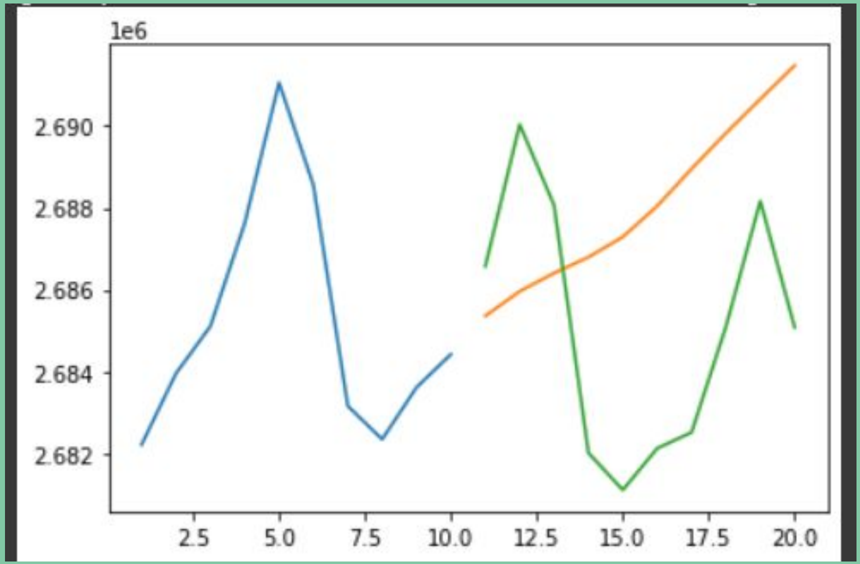
```
lst_output=[]
n_steps=10
i=0
while(i<10):

    if(len(temp_input)>10):
        #print(temp_input)
        x_input=np.array(temp_input[1:])
        print("{} day input {}".format(i,x_input))
        x_input=x_input.reshape(1,-1)
        x_input = x_input.reshape((1, n_steps, 1))
        #print(x_input)
        yhat = model.predict(x_input, verbose=0)
        print("{} day output {}".format(i,yhat))
        temp_input.extend(yhat[0].tolist())
        temp_input=temp_input[1:]
        #print(temp_input)
        lst_output.extend(yhat.tolist())
        i=i+1
    else:
        x_input = x_input.reshape((1, n_steps,1))
        yhat = model.predict(x_input, verbose=0)
        print(yhat[0])
        temp_input.extend(yhat[0].tolist())
        print(len(temp_input))
        lst_output.extend(yhat.tolist())
        i=i+1
```

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



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

The model isn't working well.
Possible reason: very few data points
In training set

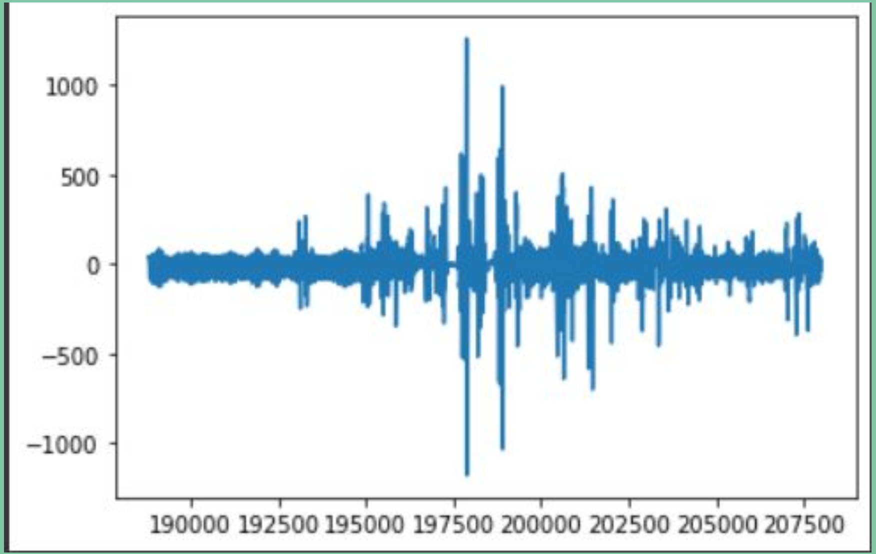


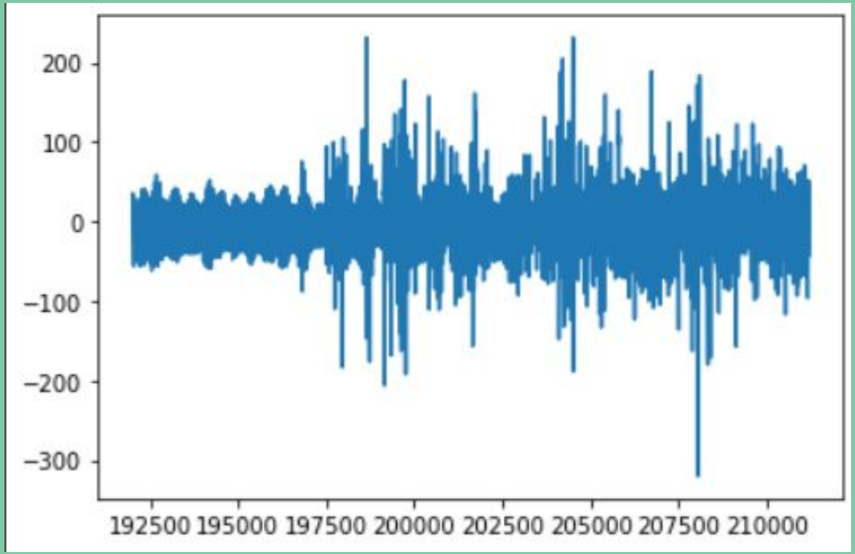
New Dataset

- Used 64Hz raw ppg data.
- Empatica E4 device
- Ground truth information obtained from the ECG-signal
- https://www.empatica.com/research/e4/
- A. Reiss, P. Schmidt, I. Indlekofer and K. V. Laerhoven. 2018. PPG-based Heart Rate Estimation with Time-Frequency Spectra: A Deep Learning Approach
- The dataset has data from 20 subjects performing various daily activities.
- Used driving activity data in the model

Updated Model

Training size: 5 minutes (64 Hz): 19200 data points

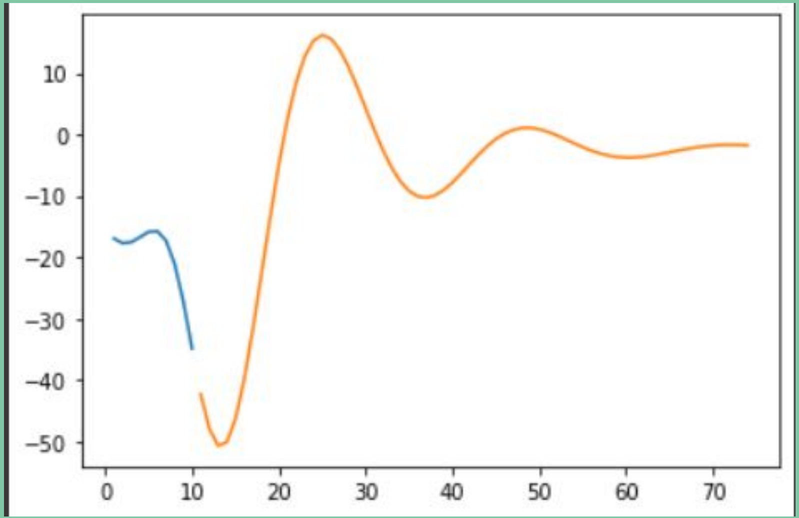


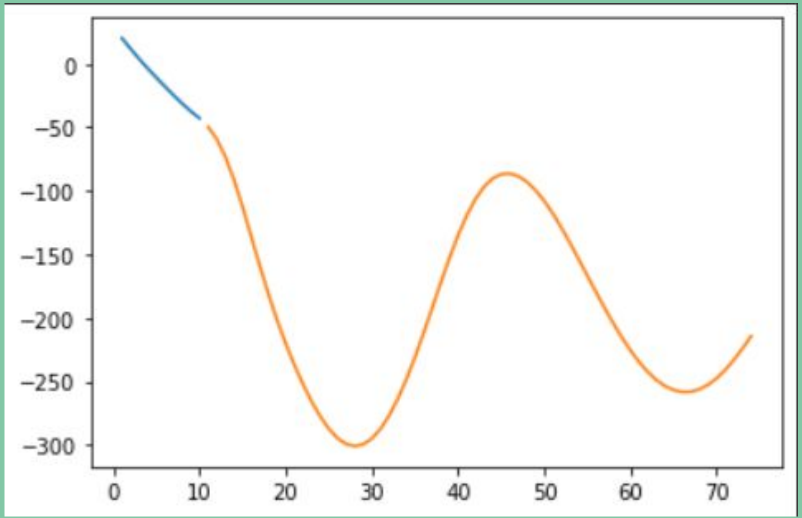


Updated Model

Training size: 5 minutes (64 Hz): 19200 data points

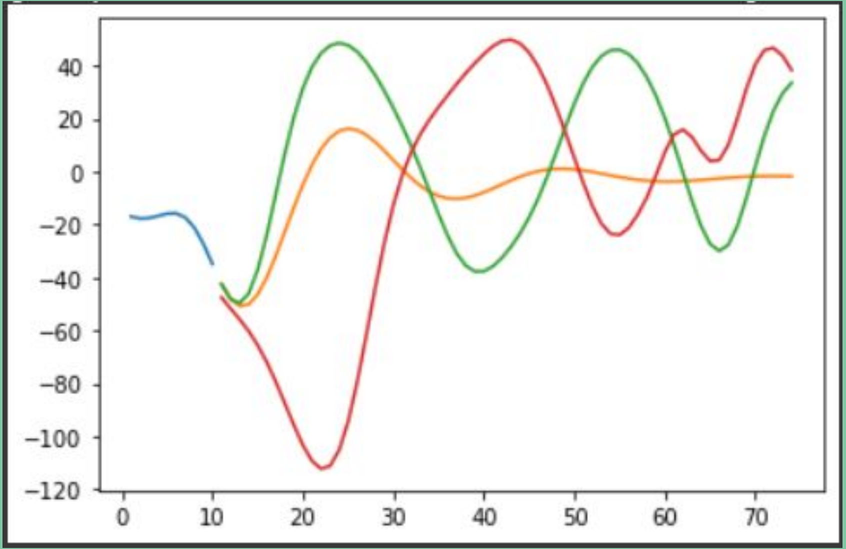
Predicted next 1 second data(64 datapoints)   Lookback: 10

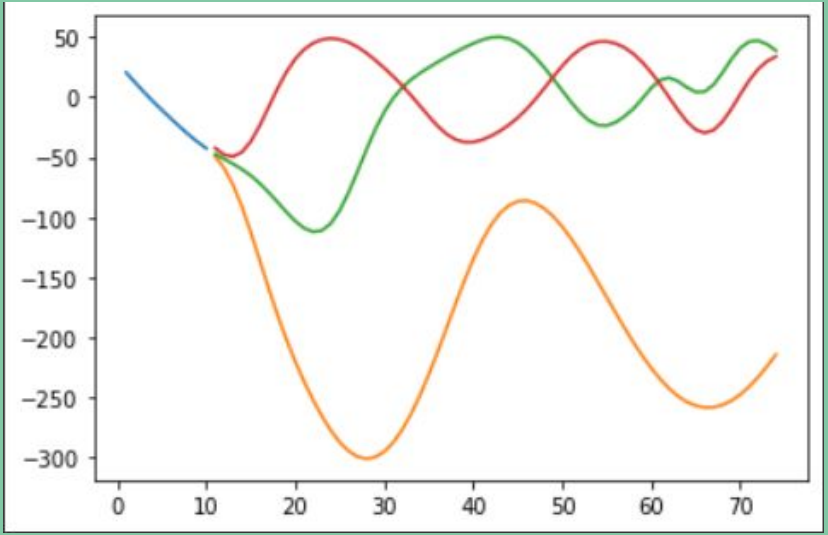




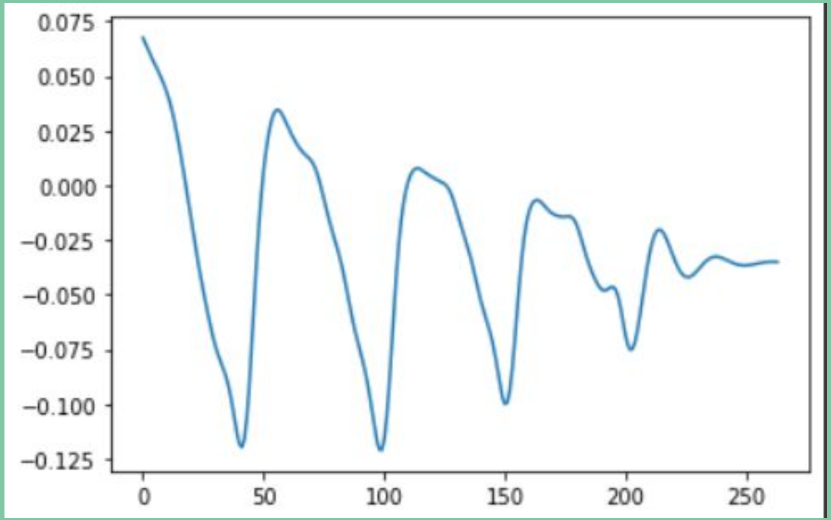
Updated Model

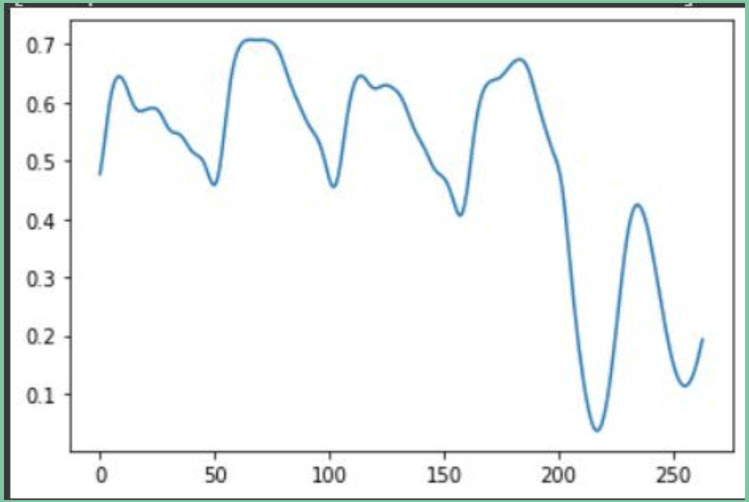
- Yellow: Predicted Data
- Green: Subject’s Real Data
- Red: Different subject data(Consciously took data with values similar to real values to have more meaningful comparison)



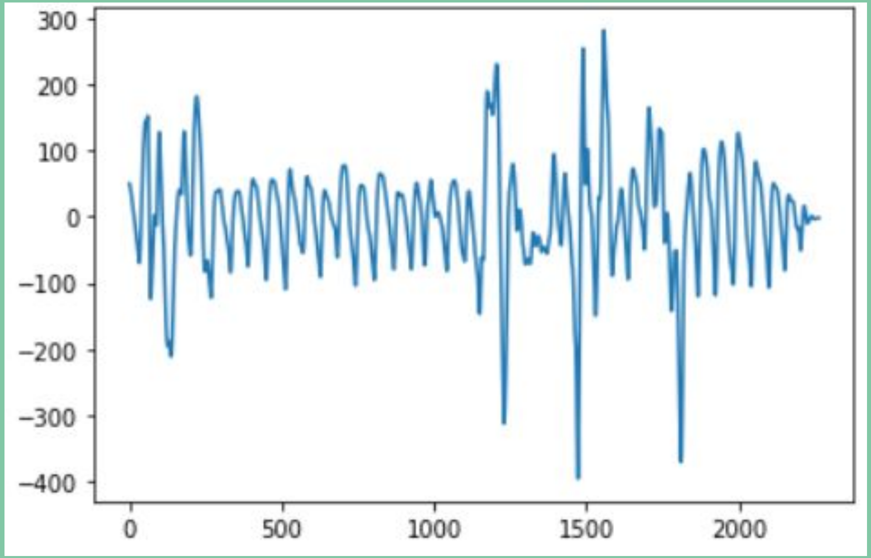


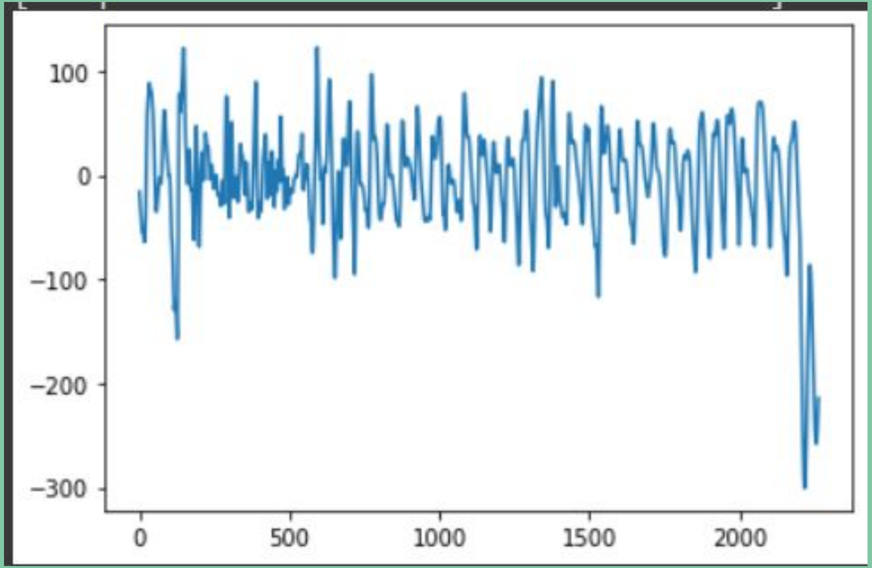
Updated Model





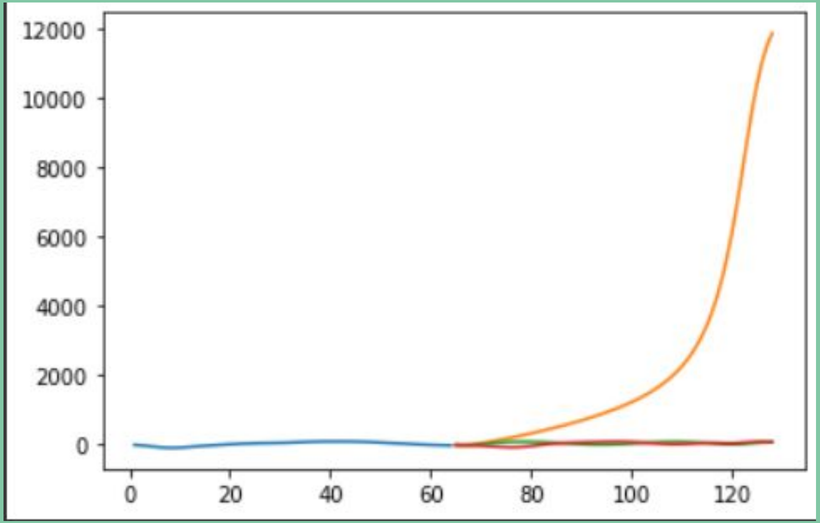
Updated Model

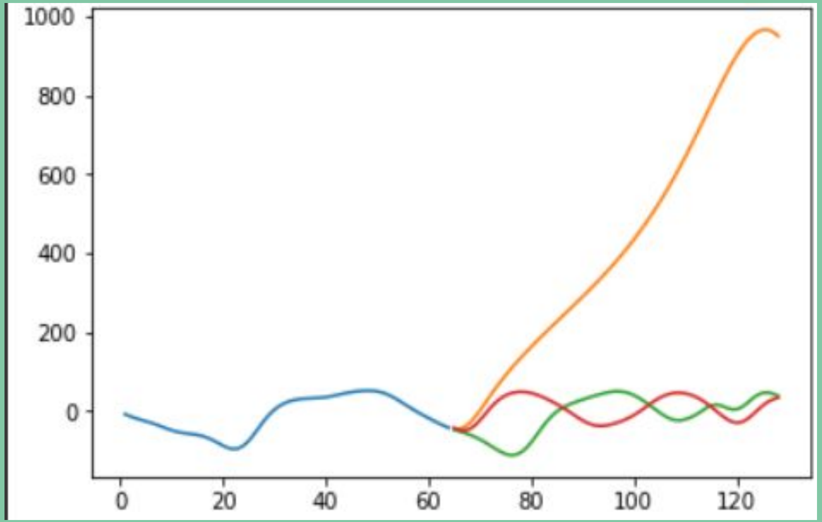




Updated Model

- Model is not working efficiently with Subject 2
- Changed lookback to 64 datapoints (1 second)
- Completely failed





Paper Implementation

- TrueHeart: Continuous Authentication on Wrist-worn Wearables Using PPG-based Biometrics
- https://ieeexplore.ieee.org/document/9155526
- Couldn’t obtain mail even after mailing the author

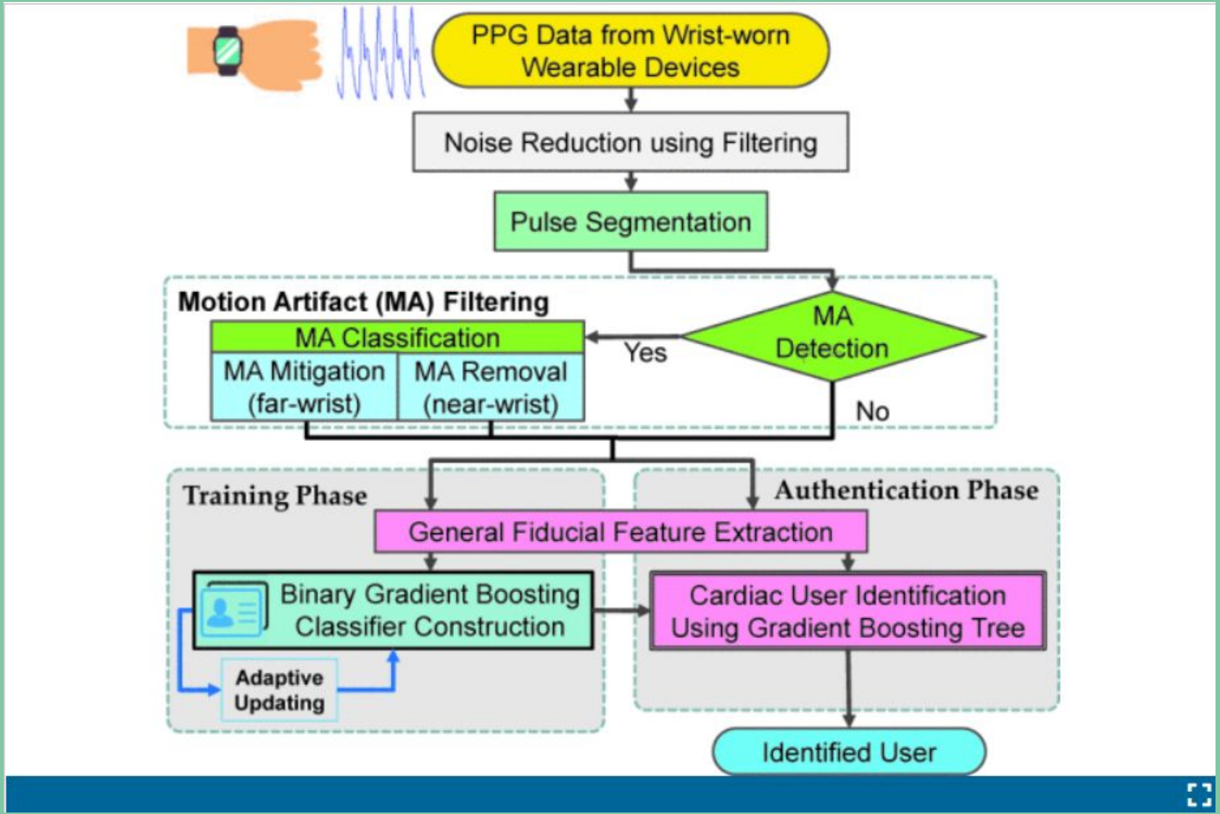
Paper Analysis

- PPG signals are relatively coarse-grained, noisy, and more susceptible to interference than ECG signals.
- Wrist-worn wearable devices are usually associated with a lot of hand or body movements from daily activities. These movements would result in various **motion artifacts (MAs)** which make cardiac signals in PPG measurements often unavailable in practice.
- Determine **general fiducial features** that are not only persistent in various users’ PPG measurements but also can capture unique characteristics of cardiac motions for CA.

Paper Analysis

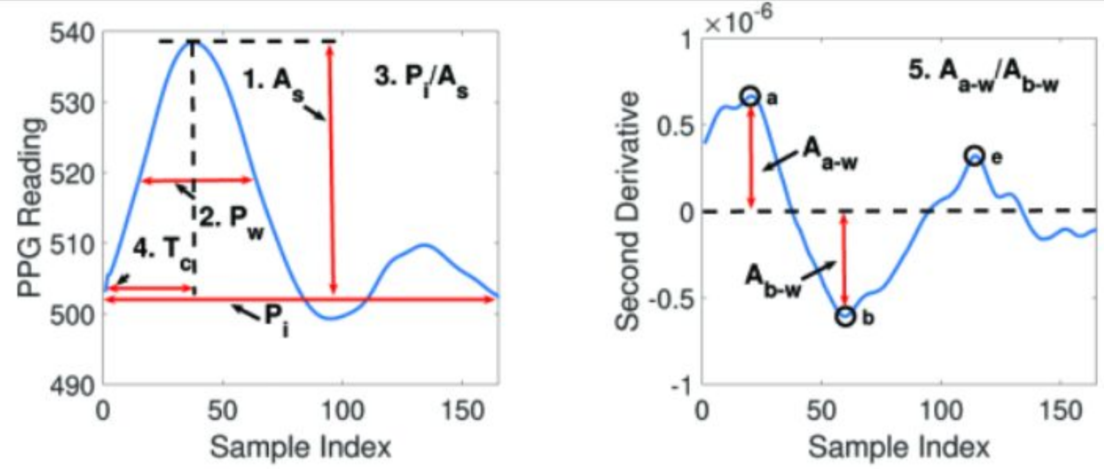
- Low-cost CA system, that can authenticate users by using unique cardiac biometrics extracted from PPG sensors in wrist-worn wearable devices. System can be easily deployed in any PPG-enabled wearable devices (e.g., smartwatches).
- Study characteristics of MAs under many practical scenarios and develop robust MA mitigation and removal mechanisms that can effectively identify different types of MAs with various intensities and eliminate MA impact accordingly.
- Identify general fiducial features that can capture the uniqueness of users' cardiac patterns to build an **adaptive gradient boosting tree (GBT)-based classifier** that can be robust to signal drifts in PPG, authenticate users, and defend against random attack effectively.

Paper Analysis



Paper Analysis

- Used five fiducial features that only require a single systolic peak in the PPG measurements



(a) Raw PPG measurements

(b) Second derivative of raw PPG measurements

Paper Analysis

Feature Name	Feature Description
Systolic Amplitude (A_s)	related to the stroke volume and directly proportional to vascular distensibility, which is distinguishable among different people.
Pulse Width (P_w)	the width of the PPG signal at the half-height of the systolic peak, and it correlates with the systemic vascular resistance.
Ratio of Pulse Interval to Systolic amplitude (P_i/A_s)	reflects the functionality of a person's cardiovascular system.
Crest Time (T_c)	indicates the pulse wave velocity, which is distinct from person to person.
Ratio of Amplitude of b-wave and a-wave (A_{b-w}/A_{a-w})	reflects the arterial stiffness and the distensibility of the peripheral artery, which are also different among people. In addition, this feature can also reflect the healthy level of different people.

Paper Analysis

- Due to hardware imperfection, the raw PPG measurements inevitably contain baseline drifts and high-frequency interferences.
- Perform **Noise Reduction using Filtering** to reduce such impacts
- A **band-pass filter** is used to extract pulsatile components in PPG measurements.
- Then the system conducts **Pulse Segmentation** to determine the **PPG segment** that is likely to contain **a complete cardiac cycle**.
- Each cardiac cycle should include a systolic peak, which could be identified in the PPG measurement during typical diastole and systole phases.
- Frequency of the pulsatile component: 0.5 – 4Hz

Paper Analysis

- Built binary classifier using Gradient Boost Tree (GBT) for user authentication
- Then **construct a binary gradient classifier bk(···) for each user gk, k = 1, ··· , K** to complete the **Training Phase**.
- Then for the testing feature set, each binary gradient classifier will output a score.
- In the **authentication phase**, system utilizes the already built binary classifiers for all the users in parallel to classify incoming cardiac-related feature set x. In particular, obtain different confidence scores from each binary classifier, and **choose the identity k of the binary classifier bk(x) with the highest score** as the final classification.

Paper Analysis

- After classification, adopt a **non-overlapped sliding window-based approach** to perform CA. Consider P (4) continuous PPG segments in a sliding window as a basic CA unit and use the majority vote from the classification results of these PPG segments to determine the user's identity periodically. If equal or more than half of the PPG segments in the window are classified to be the same user, the system would allow the current user to pass the user authentication. Otherwise, the current user does not pass the user authentication.
- **Adaptive Updating:** Re-train the underlying classifier based on the recently collected PPG measurements after each successful user authentication. Specifically, system regularly adds a small amount of the user's PPG measurements (e.g., 2min) to the training data to re-train a new classifier for the user in the background. This re-training process will stop until the new classifier meets the performance requirement (e.g., when the CA accuracy reaches 90%), and the new classifier will take effect until the next time retraining process starts.

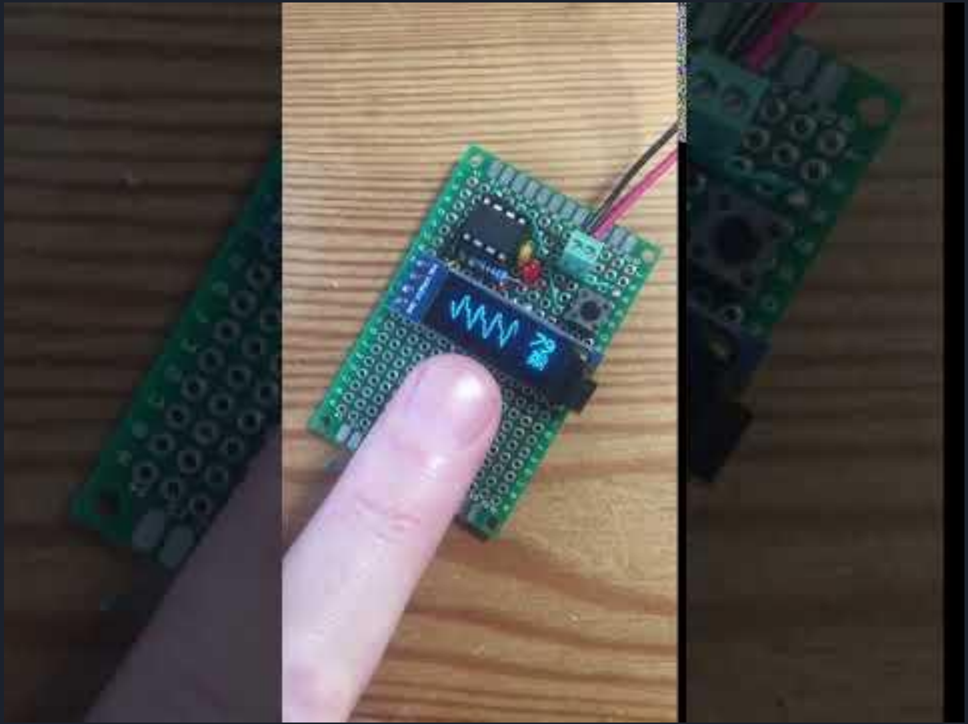
Paper Analysis

- Motion Artifacts Detection
- Motion Artifacts Classification
- Motion Artifacts Removal for Near-wrist Activities
- Motion Artifacts Mitigation for Far-wrist Activities

ATiny85 Pulse Oximeter and Photoplethysmograph

- SSD1306 128x32 OLED display
- MAX30102 sensor
- 512 Bytes RAM

Place finger screen



MAX30102 Heart Rate and Pulse Oximeter Sensor Module (Black)

Price: Rs 329 + 18%GST

Link: here

2.32 cm (0.91 inch) I2C/IIC 128x32 OLED Display Module - Blue

Price: Rs 165 + 18% GST

Link: here