4th International Conference on Electrical Information and Communication Technology (EICT), 20-22 December 2019, Khulna, Bangladesh

Extraction of Heart Rate from PPG Signal: A Machine Learning Approach using Decision Tree Regression Algorithm

Shikder Shafiul Bashar
Department of Electrical and Electronic Engineering
University of Asia Pacific
Dhaka, Bangladesh
bashar96.uap@gmail.com

Md. Sazal Miah*
Department of Electrical and Electronic Engineering
University of Asia Pacific
Dhaka, Bangladesh
sazalmiah94@gmail.com

A.H.M. Zadidul Karim
Department of Electrical and Electronic Engineering
University of Asia Pacific
Dhaka, Bangladesh
zadid@uap-bd.edu

Md. Abdullah Al Mahmud
Department of Electrical and Electronic Engineering
University of Asia Pacific
Dhaka, Bangladesh
abdullah-eee@uap-bd.edu

***Abstract*— This research article shows a novel technique which used to measure the heart rate (HR) from wearable devices such as fingertip device, wrist type device. For HR monitoring Photoplethysmography (PPG) signal is severely used. HR measurement precision is affected by noise and motion artifacts (MA) at the moment of physical movement. There are many conventional algorithms to decrease the effect of MA and measure the HR variations. In this work, a novel method called multi-model machine learning approach (MMMLA) is used to predict HR. The technique is shown in this work, which primarily trains and tests the algorithm for various features and various data sets. K-means cluster is employed for splitting noisy and non noisy data. This process helps the machine to learn from noisy and non noisy data. After separation, Decision Tree Regression method is employed to fit data and measure HR from test data. In this research feature engineering is completed in different words, a special set of the feature is chosen and check their behavior with the projected model and therefore the error rate for each set of the feature was computed. In this work, model feature is also reduced and check the behavior in estimating HR. For each case root mean square (RMS) error and average absolute error of HR was computed. The minimum average absolute error was traced in this work was 1.18 beats per minute (BPM). The outcome of this work demonstrates that the algorithm has a significant possibility to be used for PPG-based HR monitoring.**

Keywords*— *Heart Rate, Photoplethysmography, Signal processing, Machine learning, Feature engineering, Decision Tree Regression Algorithm

I. INTRODUCTION

HR observance is vital for the fitness of human body or exercisers to manage their exercising load. There is two type of technique to determine HR one is Electrocardiogram (ECG) and another one is Photoplethysmography (PPG). PPG is an easy and fewer costly optical bio-monitoring method used to compute the blood volume fluctuations which happen within the human body because of the pulsatile behavior of the vascular system [1]-[2]. Sensors placed in wrists or fingers or ear buds to collect the raw PPG signal. Motion artifact (MA) seriously affected the PPG signal during physical exercise which also affects the HR estimation. There are many false peaks because of MA that creates it troublesome to spot the particular HR-peak. There is a lot of algorithms are introduced to decrease MA from the PPG signal. These are independent component analysis (ICA) [3], wavelet denoising [4], spectral subtraction [5], empirical mode decomposition (EMD) [6], adaptive noise cancellation (ANC) [7] etc. In recent days, a novel method is introduced by Zhang et al. named TROIKA [8], and its alternative JOSS [9] to measure HR from PPG signal. Fig. 1(a) & (b) visualize a general PPG signal and periodogram of a PPG signal.

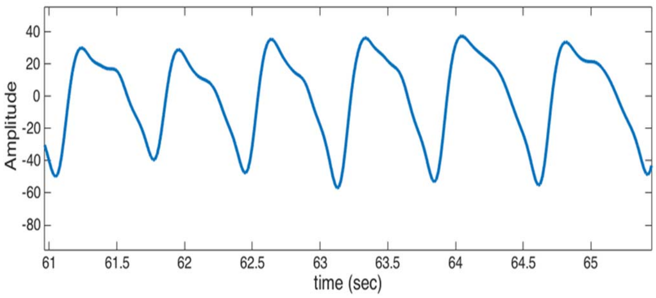

Figure 1(a): General PPG signal.

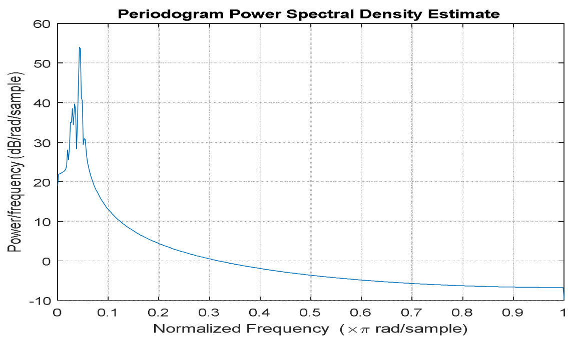

Figure- 1(b): Periodogram of a PPG signal.

Every algorithm introduced before was depended on the one sided heuristic calculation. For a one sided heuristic method, it is difficult to measure the correct HR every time for various MA. In this paper, a vigorous, computationally productive and powerful technique is proposed to calculate HR from PPG signal and acceleration data. In this work,

978-1-7281-6040-5/19/$31.00 ©2019 IEEE

different features of PPG signal also found. The core basis of this novel method is multi model machine learning. Due to AI approach now our method can distinguish HR for various and concentrated MA. In this method both noisy and non noisy data is used to train our algorithm. At first, K means clustering is used for separation of noisy and non noisy data then we introduce the Decision Tree regression algorithm which is used to fit these data and predict HR.

II. Difficulty of Traditional Frameworks & Database

A. Difficulty of Previous Traditional Algorithms

When PPG signal contains of robust MA, estimating pulse or HR becomes terribly difficult. Maximum algorithm used earlier was bias heuristic. In a one sided heuristic algorithm, it has same approach for every MA corrupted PPG signal, however MA isn't same every time. So, it's not always possible to estimate accurate HR. That is why a novel technique is applied here to estimate HR from PPG signal. Here the core basis of our approach is multi-model machine learning. In this method, the machine learning process is done here by using two algorithms at the same time.

B. Data documentation

Every datasets applied in this work were initially recorded and utilized in [8]. For the study, twelve volunteers (male, with ages starting from 18 to 35) were enrolled. PPG signals, a single-channel ECG signal and acceleration signals were recorded for every volunteer during the physical activities. Then the collected dataset were sampled at 125 Hz. For collecting PPG signal from wrist a pulse oximeter is used which contains green LED with a wavelength of 515 nm. Both the pulse oximeter and the accelerometer were integrated in a wristband, adequately worn on the subject's radiocarpal joint or wrist. For the collection of ECG signal, wet ECG sensors were placed on chest which also gives the ground-true HR.

III. Methodology

At first raw PPG signal and acceleration signal is collected synchronously. Then 2nd order bandpass filter whose frequency range starting from 0.4 to 5 Hz used for filtration process of PPG signal. In this work, machine learning is applied to estimate HR. In this research, the machine learns from every noisy and non noisy data and detects the HR. In our proposed algorithm, K means clustering and Decision Tree regression framework are used. Firstly, all algorithm features are applied to train the algorithm and then the algorithm features are reduced and checked the behavior in determining HR.

A. Primary processing of data

For primary processing of data an adaptive filter is used. The PPG signal and acceleration signal preprocessed here. The main purpose of this primary preprocessing is to remove noise and MA outside of the frequency band. Then these preprocessed signals were standardized to zero mean and unit variance.

B. Feature Engineering

After Spectral Subtraction we have a tendency to try to spot the feature of PPG signal which is employed to observe pulse or HR. The dominant feature that is employed to observe HR such as Peak position after Spectral Subtraction of first PPG signal and acceleration signal, Peak position of first PPG signal and second PPG signal, Acceleration signal's peak power and position, PPG signal one and two peak power, X, Y, Z-axis acceleration signal's peak power and position. Firstly, every feature set are applied to train and test our algorithm for the various data set. During this work, seven completely different set of feature additionally created for K means clustering. Then the various feature set is employed to cluster the data and check the response of each feature set in estimating HR.

C. K means Clustering

The process to gather information that have similar characteristics into sets of groups is termed as agglomeration or clustering. One of the clustering techniques is K means clustering. K means clustering builds every dataset in a dimensional area into K clusters [10]. Cluster means ($m^{(k)}$) is used to specify each cluster. The process of K means clustering method is shown here. At first mark K means m^k which present centroids, Then repeat the same process again. After that every datasets were appointed to the closest centroids in the form K clusters.

$$\mathrm{k(n)} = argmin_x \mathrm{d}\left(m^{(k)}, x^{(n)}\right) \tag{1}$$

Then every authorized datasets have been used to re-calculate the centroid of each cluster.

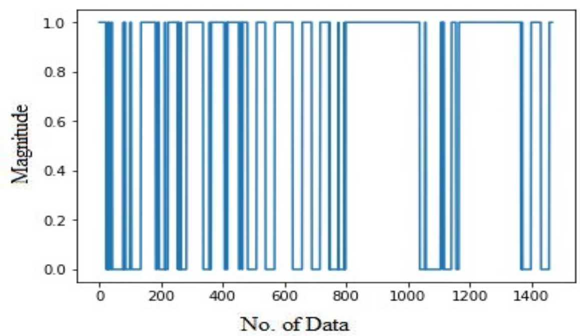

Figure- 2(a): Visualization of train data after K means clustering.

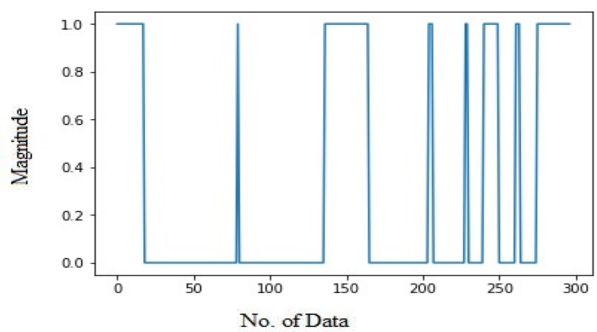

Figure- 2(b): Visualization of test data after K means clustering.

Fig. 2 visualizes the output of K means clustering where noisy and non noisy data is separated by 0 and 1 respectively.

D. Decision Tree Regression

For both classification and regression problems, Decision Tree Regression algorithm is used. The decision tree classification pseudocode is shown below:

1. Place the most effective attribute of the data set within the root of the tree.

2. Split the training set into sub-sets. And every sub-set must contain information with the similar value for an attribute.

3. Repeat both of previous steps on every sub-set till you discover leaf nodes in every branches of the tree.

The decision tree prediction pseudocode is shown below:

1. Initially, the entire training set is treated as the root.

2. Feature values are chosen to be categorical.

3. On the idea of attribute values, documentations are allocated recursively.

4. Order to inserting attributes as root or inner node of the tree is completed by applying statistical method.

E. Heart Rate detection

In this research, clustered train information and actual heart rate is accustomed to fit in Decision Tree regression framework for every noisy and non noisy datasets. So, that the algorithm will learn by it-self and may estimate the heart rate. During this research, test datasets are accustomed estimate the heart rate by applying the trained algorithm. During this analysis, heart rate is predicted for every noisy and non noisy datasets.

IV. Comparison of Proposed Algorithm Result with Ground-True HR

In this analysis, every algorithm features are applied to train the framework. Here, the comparability of two type of datasets are visualized, first one is where ground-true HR is almost matched with the calculated HR. Another one is where ground-true HR is mismatched with calculated HR for almost every datasets.

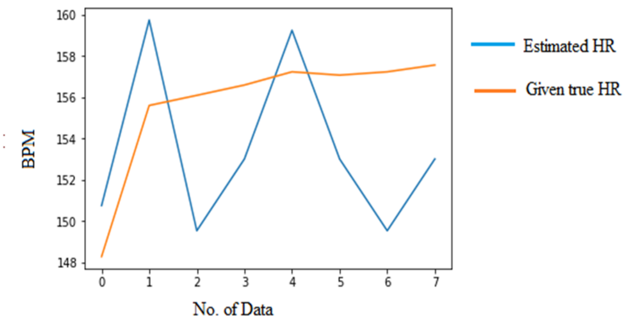

Figure- 3(a): Visualization of comparability between calculated and ground-true HR of noisy data for matched case when Decision Tree Regression applied with every algorithm feature.

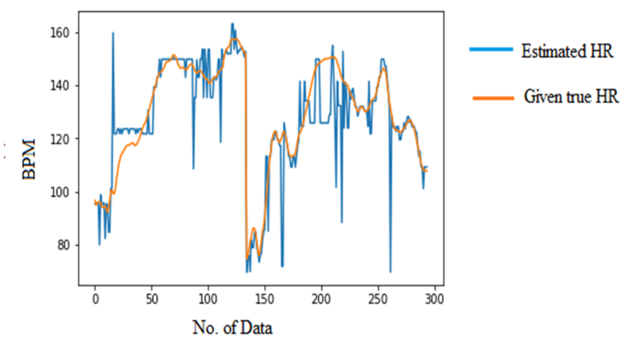

Figure- 3(b): Visualization of comparability between calculated and ground-true HR of non noisy data for matched case when Decision Tree Regression applied with every algorithm feature.

In this research, matched case occurred for feature set which is introduced as Peak position after Spectral Subtraction of first PPG signal and acceleration signal (sspeak1) and acceleration signal & first PPG signal peak power (ppg1). For train the model the datasets used here are 1-6, 9-12 and dataset 7, 8 are used here for testing purpose which visualized in Fig. 3(a) & (b).

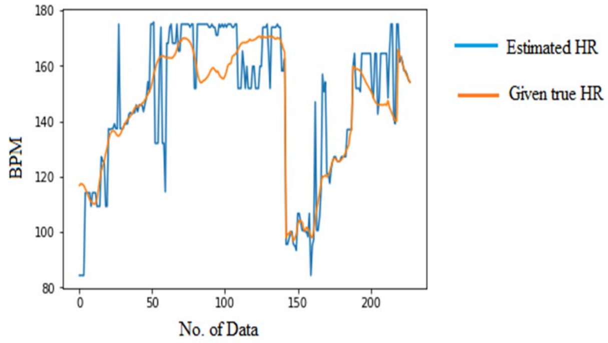

Figure- 4(a): Visualization of comparability between calculated and ground-true HR of noisy data for mismatched case when Decision Tree Regression applied with every algorithm feature.

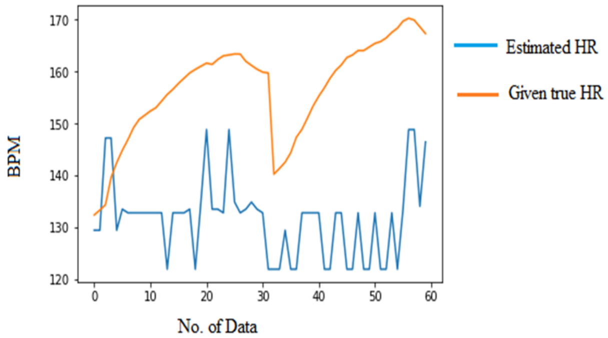

Figure- 4(b): Visualization of comparability between calculated and ground-true HR of non noisy data for mismatched case when Decision Tree Regression applied with every algorithm feature.

In this work, mismatched case occurred for feature set which is introduced as Acceleration signal's peak power. For train the model dataset 1-10 used here and dataset 11, 12 used here for testing purpose visualized in Fig. 4(a) & (b).

Now, the model feature is reduced to find the behavior of our developed model and check whether a higher number of the model feature is good for our developed model or not. Here, the matched case occurred for feature set which is introduced as Peak position after Spectral Subtraction of first PPG signal and acceleration signal (sspeak1) and acceleration signal & first PPG signal peak power (ppg1). For train the model the datasets used here are 1-6, 9-12 and dataset 7, 8 are used here for testing purpose which visualized in Fig. 5(a) & (b).

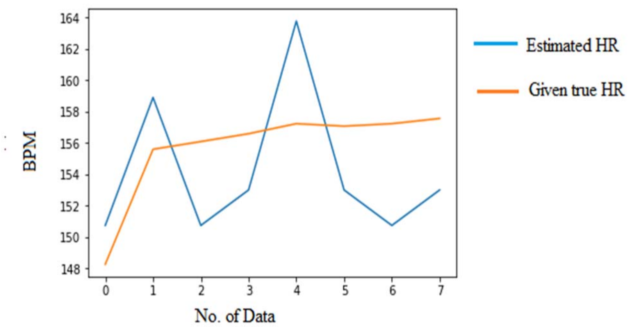

Figure- 5(a): Visualization of comparability between calculated and ground-true HR of noisy data for matched case when Decision Tree Regression applied with fewer algorithm feature.

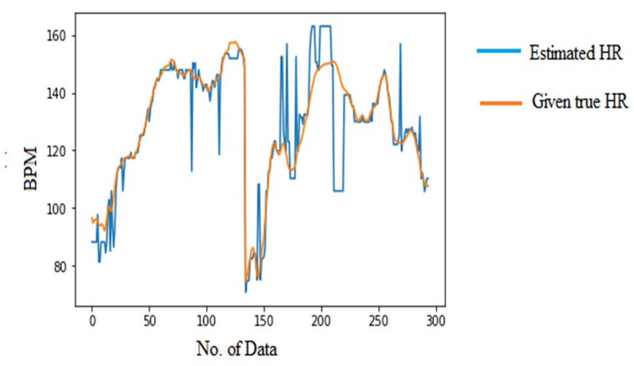

Figure- 5(b): Visualization of comparability between calculated and ground-true HR of non noisy data for matched case when Decision Tree Regression applied with fewer algorithm feature.

In this case, the clustering feature set for the matched case is the same as before when all model features are taken. From this analysis, we reached to this decision that clustering feature named Spectral Subtraction of first PPG signal and acceleration signal (sspeak1) and acceleration signal & first PPG signal peak power (ppg1) is good for training the data.

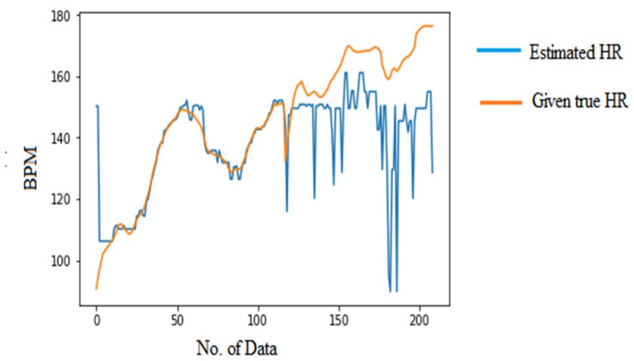

Figure- 6(a): Visualization of comparability between calculated and ground-true HR of noisy data for mismatched case when Decision Tree Regression applied with fewer algorithm feature.

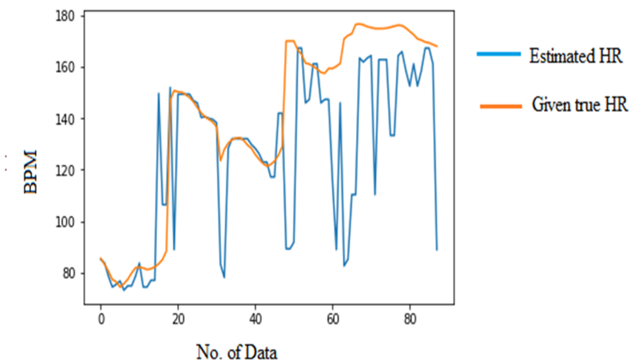

Figure- 6(b): Visualization of comparability between calculated and ground-true HR of non noisy data for mismatched case when Decision Tree Regression applied with fewer algorithm feature.

In this analysis, the mismatched case occurred for feature set which is introduced as Acceleration signal's peak power. For train the model the datasets used here are 1-8, 11, 12 and dataset 9, 10 are used here for testing purpose which visualized in Fig. 6(a) & (b). The clustering feature set for the mismatched case is the same as before when all model features are taken but the data set is different. Now we reached to this decision that, this clustering feature is not a good option for train data. And when smaller numbers of model features are taken, than it takes lower time to train data and the result is almost the same when all model features are taken sometimes better. Now we can say that the lower model feature is good for our developed model.

V. ERROR CALCULATION

For determination of performance of the proposed algorithm, we determine the Average absolute error denoted as,

$$A_{error} = \frac{1}{N}\sum_{i=1}^{N} |HR_{est(i)} - HR_{true(i)}| \qquad (2)$$

And we also calculate RMS error denoted as,

$$\mathrm{RMS}_{error} = \sqrt{\frac{1}{N}\left(\sum_{i=1}^{N} HR_{est(i)} - HR_{true(i)}\right)^2} \qquad (3)$$

Here, ground-truth heart rate denoted as $HR_{true(i)}$ and calculated heart rate denoted as $HR_{est(i)}$. And total number of time windows denoted as N. For both noisy & non noisy signal, the average absolute error and RMS error was measured for each data-set and each clustering feature set. Average absolute error of noisy signal is denoted as abs_err0 and rms0 used to denote RMS error of noisy signal. And average absolute error of non noisy signal is denoted as abs_err1 and rms1 used to denote RMS error of non noisy signal. In this work, both errors are calculated in Beats per Minute (BPM).

Table- I: Error calculation for Acceleration peak power and each dataset for Decision Tree algorithm.

Clustering Feature Name: Acceleration peak power					
Algorithm Name	Test Data	abs_err0	abs_err1	rms0	rms1
Decision Tree For all model feature	1,2	10.34	11.32	19.61	16.28
	3,4	11.30	7.64	20.20	13.10
	5,6	7.28	15.58	12.52	22.13
	7,8	9.75	6.68	15.89	12.59
	9,10	10.71	12.75	18.16	24.45
	11,12	8.73	24.75	12.25	27.27
Decision Tree For fewer model feature	1,2	10.59	18.97	19.43	26.27
	3,4	10.69	8.25	19.55	13.40
	5,6	7.41	12.19	13.72	18.21
	7,8	9.75	5.87	15.87	11.44
	9,10	9.72	16.15	16.82	28.74
	11,12	13.95	9.37	20.32	10.69

Table- II: Error calculation for ppg1 & acceleration peak position and each dataset for Decision Tree algorithm.

Clustering Feature Name: ppg1 & acceleration peak position					
Algorithm Name	Test Data	abs_err0	abs_err1	rms0	rms1
Decision Tree For all model feature	1,2	11.54	1.18	18.40	1.44
	3,4	8.24	2.01	16.38	2.82
	5,6	7.12	4.92	13.59	5.11
	7,8	6.71	4.48	13.11	4.87
	9,10	16.64	3.13	26.47	5.21
	11,12	5.29	8.66	7.87	12.57
Decision Tree For fewer model feature	1,2	12.63	2.15	20.77	2.53
	3,4	11.20	2.26	20.16	2.83
	5,6	4.42	5.17	10.19	5.19
	7,8	5.97	4.31	11.55	4.52
	9,10	17.23	3.94	27.25	5.97
	11,12	5.02	7.20	7.93	11.28

Table- III: Error calculation for ppg1 & acceleration peak power and each dataset for Decision Tree algorithm.

Clustering Feature Name: ppg1 & acceleration peak power					
Algorithm Name	Test Data	abs_err0	abs_err1	rms0	rms1
Decision Tree For all model feature	1,2	11.99	1.47	17.72	1.71
	3,4	7.30	2.35	14.98	3.01
	5,6	7.05	4.92	13.45	5.11
	7,8	6.15	5.10	11.93	5.90
	9,10	14.41	3.23	21.86	5.33
	11,12	9.53	5.58	14.14	8.34
Decision Tree For fewer model feature	1,2	11.96	1.89	19.98	2.34
	3,4	11.13	2.41	20.09	2.92
	5,6	4.67	5.49	10.13	5.55
	7,8	5.01	3.98	10.31	4.21
	9,10	14.96	3.71	23.55	5.31
	11,12	7.12	7.77	11.34	11.66

Table- IV: Error calculation for ppg2 & acceleration peak power and each dataset for Decision Tree algorithm.

Clustering Feature Name: ppg2 & acceleration peak power					
Algorithm Name	Test Data	abs_err0	abs_err1	rms0	rms1
Decision Tree For all model feature	1,2	12.84	12.12	16.96	18.97
	3,4	6.61	N/A	14.15	N/A
	5,6	8.07	N/A	15.40	N/A
	7,8	5.31	N/A	11.10	N/A
	9,10	13.11	14.75	22.52	19.46
	11,12	13.99	6.35	18.59	11.65
Decision Tree For fewer model feature	1,2	10.15	10.26	14.81	18.06
	3,4	10.26	N/A	13.34	N/A
	5,6	10.26	N/A	8.54	N/A
	7,8	10.26	N/A	10.13	N/A
	9,10	12.18	7.37	22.32	10.00
	11,12	13.20	5.61	19.22	10.41

Table- V: Error calculation for sspeak1 and each dataset for Decision Tree algorithm.

Clustering Feature Name: sspeak1					
Algorithm Name	Test Data	abs_err0	abs_err1	rms0	rms1
Decision Tree For all model feature	1,2	12.00	5.92	17.97	11.96
	3,4	6.82	4.85	14.36	11.02
	5,6	2.46	9.56	6.17	16.01
	7,8	5.39	2.15	9.60	7.55
	9,10	6.21	17.81	15.79	25.00
	11,12	6.59	14.90	14.18	20.91
Decision Tree For fewer model feature	1,2	15.05	5.92	21.70	12.01
	3,4	8.41	4.91	15.29	12.73
	5,6	3.65	12.17	10.40	19.58
	7,8	4.90	2.43	10.34	8.47
	9,10	5.61	12.64	15.95	17.76
	11,12	5.73	9.01	11.33	13.01

Table- VI: Error calculation for sspeak1 & ppg1 and each dataset for Decision Tree algorithm.

Clustering Feature Name: sspeak1 & ppg1					
Algorithm Name	Test Data	abs_err0	abs_err1	rms0	rms1
Decision Tree For all model feature	1,2	11.32	1.67	17.44	2.02
	3,4	8.31	6.59	15.39	12.69
	5,6	6.85	3.49	13.11	3.62
	7,8	3.41	6.64	4.18	12.67
	9,10	13.89	3.57	21.76	4.90
	11,12	4.88	9.89	7.67	14.71
Decision Tree For fewer model feature	1,2	11.90	2.17	19.50	2.49
	3,4	11.97	1.91	21.32	2.47
	5,6	5.91	3.49	13.79	3.62
	7,8	2.49	5.16	2.90	10.35
	9,10	17.03	4.02	28.13	6.11
	11,12	6.12	7.11	8.95	11.03

Table- VII: Error calculation for sspeak1 & ppg2 and each dataset for Decision Tree algorithm.

Clustering Feature Name: sspeak1 & ppg2					
Algorithm Name	Test Data	abs_err0	abs_err1	rms0	rms1
Decision Tree For all model feature	1,2	12.70	12.96	19.34	16.80
	3,4	6.13	N/A	13.82	N/A
	5,6	7.82	N/A	15.85	N/A
	7,8	4.41	N/A	7.97	N/A
	9,10	12.91	15.43	22.37	19.90
	11,12	12.23	6.53	17.37	11.62
Decision Tree For fewer model feature	1,2	10.67	8.53	18.33	13.96
	3,4	8.82	N/A	14.01	N/A
	5,6	3.47	N/A	8.56	N/A
	7,8	3.45	N/A	8.49	N/A
	9,10	11.84	7.69	22.05	10.76
	11,12	13.08	5.72	18.69	11.51

Form this analysis we reached to these decisions that, clustering feature introduced as Peak position after Spectral Subtraction of first PPG signal and acceleration signal (sspeak1) has a low error rate in almost each noisy cases. First PPG signal peak power (ppg1) & Acceleration signal's peak position has low error rate for each dataset and we will get the better prediction of HR when we will train our model with these feature set for non noisy data. And we get low error rate for test dataset 7, 8 for each clustering feature. This also applies for when we take lower number of model features to train our model. For better understanding, the plot for Peak position after Spectral Subtraction of first PPG signal and acceleration signal (sspeak1) are shown below.

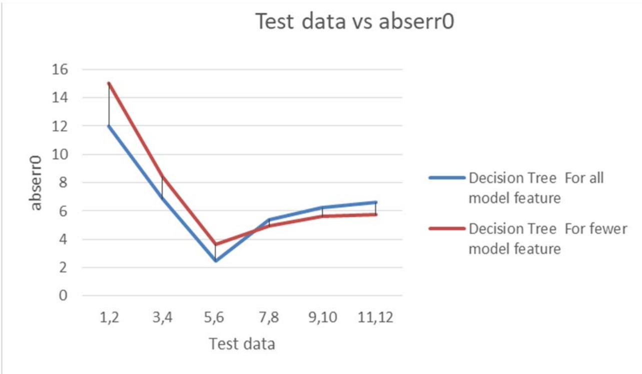

Figure- 7(a): Graphical representation of each test data with average absolute error for noisy data with Decision Tree algorithm for sspeak1.

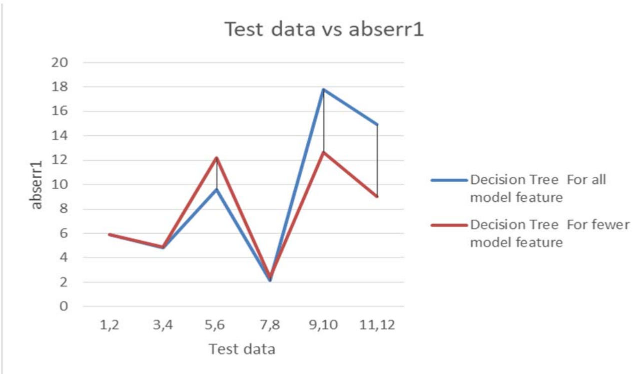

Figure- 7(b): Graphical representation of each test data with average absolute error for non noisy data with Decision Tree algorithm for sspeak1.

In Fig. 7 (a) & (b) demonstrates the comparability of average absolute error rate when we take all features and fewer model features for sspeak1. From figure we can see that for maximum test data set the absolute error rate is reduced or same when we take fewer model features.

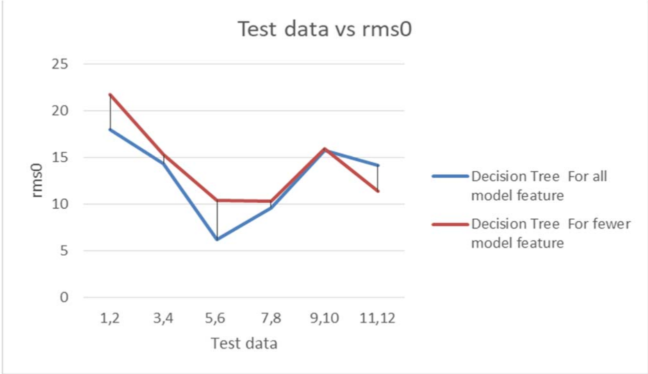

Figure- 8(a): Graphical representation of each test data with root mean square error for noisy data with Decision Tree algorithm for sspeak1.

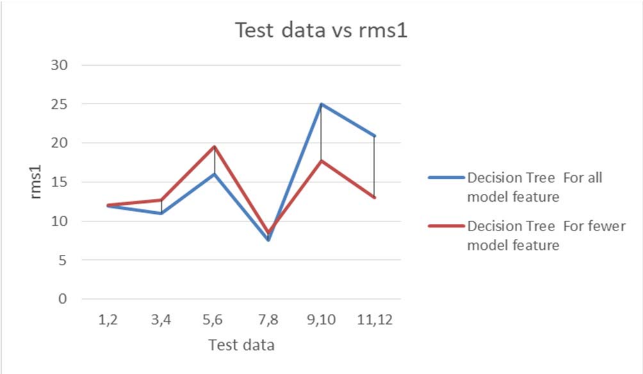

Figure- 8(b): Graphical representation of each test data with root mean square error for non noisy data with Decision Tree algorithm for sspeak1.

In Fig. 8 (a) & (b) shows the comparison of root mean square error rate when we take all features and fewer model features for sspeak1. From figure we can see that for maximum test data set the absolute error rate is reduced or same when we take fewer model features. As we found, the time required to train the model with fewer features is less then when we take all model features. That's why we can say that, fewer model features for training purpose is good for our model.

VI. Conclusion

In this research, multi-model machine learning approach (MMMLA) is applied to predict heart rate. In this research, the feature engineering is also done. In addition; different set of feature is used here to find their characteristics in training data and calculating heart rate. For separation of noisy and non noisy data K means clustering is used here. And then these separated data used to fit into Decision Tree regression framework to calculate heart rate. In this work, the model features are also reduced and check their behavior in determining heart rate.

Acknowledgment

The research was conducted in the Department of Electrical and Electronic Engineering, University of Asia Pacific, Dhaka-1215, Bangladesh.

References

[1] Allen, John. "Photoplethysmography and its application in clinical physiological measurement." Physiological measurement 28, no. 3 (2007): R1.

[2] Bashar, Shikder Shafiul, Md Sazal Miah, AHM Zadidul Karim, Md Abdullah Al Mahmud, and Zahid Hasan. "A Machine Learning Approach for Heart Rate Estimation from PPG Signal using Random Forest Regression Algorithm." In 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1-5. IEEE, 2019.

[3] Kim, Byung S., and Sun K. Yoo. "Motion artifact reduction in photoplethysmography using independent component analysis." IEEE transactions on biomedical engineering 53, no. 3 (2006): 566-568.

[4] Raghuram, M., K. Venu Madhav, E. Hari Krishna, and K. Ashoka Reddy. "Evaluation of wavelets for reduction of motion artifacts in photoplethysmographic signals." In 10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010), pp. 460-463. IEEE, 2010.

[5] Fukushima, Hayato, Haruki Kawanaka, Md Shoaib Bhuiyan, and Koji Oguri. "Estimating heart rate using wrist-type photoplethysmography and acceleration sensor while running." In 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2901-2904. IEEE, 2012.

[6] Sun, Xuxue, Ping Yang, Yulin Li, Zhifan Gao, and Yuan-Ting Zhang. "Robust heart beat detection from photoplethysmography interlaced with motion artifacts based on empirical mode decomposition." In Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics, pp. 775-778. IEEE, 2012.

[7] Ram, M. Raghu, K. Venu Madhav, E. Hari Krishna, Nagarjuna Reddy Komalla, and K. Ashoka Reddy. "A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter." IEEE Transactions on Instrumentation and Measurement 61, no. 5 (2011): 1445-1457.

[8] Zhang, Zhilin, Zhouyue Pi, and Benyuan Liu. "TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise." IEEE Transactions on biomedical engineering 62, no. 2 (2014): 522-531.

[9] Zhang, Zhilin. "Photoplethysmography-based heart rate monitoring in physical activities via joint sparse spectrum reconstruction." IEEE transactions on biomedical engineering 62, no. 8 (2015): 1902-1910.

[10] Djawad, Yasser Abd, Andi Mu'nisa, Pangayoman Rusung, Abdi Kurniawan, Irma Suryani Idris, and Mushawwir Taiyeb. "Essential Feature Extraction of Photoplethysmography Signal of Men and Women in Their 20s." Engineering Journal 21, no. 4 (2017): 259-272.