# Blood Pressure Estimation using Photoplethysmography only: Comparison between Different Machine Learning Approaches

Syed Ghufran Khalid[1*], Jufen Zhang[1], Fei Chen[2], Dingchang Zheng[1*]

[1]Faculty of Medical Science, Anglia Ruskin University, Bishop Hall Ln, Chelmsford, CM11SQ, United Kingdom

[2]Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen 518055, China

*: Corresponding Authors:

Syed Ghufran Khalid and Dingchang Zheng

Faculty of Medical Science,

Anglia Ruskin University,

Bishop Hall Ln,

Chelmsford, CM11SQ

United Kingdom

Syed.Khalid1@pgr.anglia.ac.uk

Dingchang.zheng@anglia.ac.uk

# Abstract

## Introduction

Blood pressure (BP) has been a potential risk factor for cardiovascular diseases. BP measurement is one of the useful parameters for early diagnosis, prevention, and treatment of cardiovascular diseases. At present, BP measurement mainly relies on cuff-based techniques that cause inconvenience and discomfort to users. Although some of the present prototype cuffless BP measurement techniques are able to reach overall acceptable accuracies, they require an electrocardiogram (ECG) and photoplethysmograph (PPG) that makes them unsuitable for true wearable applications. Therefore, developing a single PPG based cuffless BP estimation algorithm with enough accuracy would be clinically and practically useful.

## Methods

The University of Queensland vital sign dataset (Online database) was accessed to extract raw PPG signals and its corresponding reference BPs (Systolic BP & Diastolic BP). The online database consisted of PPG waveforms of 32 cases from whom 8133 (good quality) signal segments (5s for each) were extracted, pre-processed and normalised in both width and amplitude. Three most significant features (Pulse area, Pulse Rising Time and Width 25%) with their corresponding reference BPs were used to train and test three machine learning algorithms (Regression Tree, Multiple Linear Regression (MLR) and Support Vector Machine (SVM)). A 10-fold cross-validation was applied to obtain over-all BP estimation accuracy, separately for the three machine learning algorithms. Their estimation accuracies were further analysed separately for three clinical BP categories (Normotensive, Hypertensive and Hypotensive). Finally, they were compared with the ISO standard for non-invasive BP device validation (average difference no greater than 5mmHg and SD no greater than 8mmHg).

## Results

In terms of overall measurement accuracy, the Regression Tree achieved the best overall accuracy for SBP (mean and SD of difference: -0.1±6.5mmHg) & DBP (mean and SD of difference: -0.6±5.2mmHg). MLR and SVM achieved the overall mean difference less than 5mmHg for both SBP and DBP but their SD of difference was >8mmHg. Regarding the measurement accuracy in each BP categories, only the Regression Tree achieved acceptable ISO standard for SBP (-1.1±5.7mmHg) & DBP (-0.03±5.6 mmHg) in the Normotensive category. MLR and SVM did not achieve acceptable accuracies in any BP categories.

## Conclusion

This study developed and compared three machine learning algorithms to estimate BPs using PPG only, and revealed that the Regression Tree algorithm was the best approach with overall acceptable accuracy to ISO standard for BP device validation. Furthermore, this study demonstrated that the Regression Tree algorithm achieved acceptable measurement accuracy only in the Normotensive category, suggesting that future algorithm development for BP estimation should be more specific for different BP categories.

## Introduction

Blood Pressure (BP) is one of the main risk factors for cardiovascular diseases. Abnormal BP has been a potent issue that causes strokes, heart attacks and kidney failure (Höcht 2013). At present, cuff-based BP measurement devices have been widely used in hospital settings to detect abnormal BP (de la Sierra 2017). However, they are not convenient and comfortable for the users.

In the past few years, various research groups have attempted numerous techniques in order to achieve cuffless BP measurement. The key measurement principle for cuffless BP estimation is based upon the time taken by a pulse from the heart to the finger. They are known as Pulse transit time (PTT) or Pulse arrival time (PAT) (Buxi, Redoute et al. 2015, Ding, Zhang et al. 2016, Ding, Yan et al. 2017, Wong, Pickwell-MacPherson et al. 2009, Wan-Hua, Hui et al. 2017, Gesche, Grosskurth et al. 2012, Chen, Kobayashi et al. 2000, Kachuee, Kiani et al. 2017). Other researchers used vascular transit time (VTT) which was calculated from the time difference between PPG measured at fingertip and phonocardiograph measured at the chest (Shukla, Kakwani et al. 2015). Cuffless BPs were also measured using tonometry technique based on the information from multiple pressure sensors on the radial artery tree (Ding, Zhang et al. 2016, Park, Kang et al. 2007). Another group of researchers introduced cuffless BP measurement technique using modified normalized pulse volume and heart rate (Matsumura, K, et al. 2018). Multiple magnetic sensors have also been used to measure pulse wave velocity (PWV) for the estimation of cuffless BP (Nabeel, Joseph et al. 2014). Although some of the cuffless BP devices achieved overall acceptable accuracies, the above mentioned algorithms required at least two sensors (McCarthy, Vaughan et al. 2013) making them unsuitable for true wearable applications. Therefore, developing a single photoplethysmograph (PPG) based cuffless BP estimation algorithm with enough accuracy would be clinically and practically useful.
Recently, machine learning algorithms, including support vector machine (SVM), multiple linear regression (MLR) and Neural Networks algorithms have been used to estimate cuffless BP. Zhang and Feng applied SVM algorithm to waveform features that were extracted from PPG signal segments collected from the University of Queensland Vital Signs dataset (Zhang Y and Feng Z, 2017). Nevertheless, their study only achieved the SBP and DBP measurement accuracies of 11.6 ± 8.2 mmHg & 7.6 ± 6.7 mmHg (Zhang Y and Feng Z, 2017). Kawanaka et al tested linear regression with their own collected dataset. Their training data involved old individuals while testing datasets gathered from young individuals (Atomi, Kawanaka et al. 2017). Visnathana et al also used PPG signal features with both linear regression and SVM algorithms to estimate cuffless BP (Visvanathan, Sinha et al. 2013). However, these studies failed to meet ISO non-invasive BP device accuracy (average difference no greater than 5mmHg and SD no greater than 8mmHg). Other researchers also developed a cuffless BP measurement device with acceptable in terms of mean difference (3.8 mmHg for SBP and 4.6 mmHg for DBP) accuracy, but unfortunately, their measurement technique has not described in details (Watanabe, Bando et al. 2017). Furthermore, in all the published studies, the measurement accuracies have not been evaluated specifically in different clinical BP categories (Normotensive, Hypertensive and Hypotensive).

This research aimed to develop and compare three machine learning algorithms (Regression Tree, MLR, and SVM) to estimate BPs only using pulse waveform features derived from good quality PPG signals. In addition, their measurement accuracy would be evaluated in three different clinical BP categories (Normotensive, Hypertensive and Hypotensive).

## Methods

The overall flow diagram of the proposed research methodology is presented in Figure 1, which is summarised in the following steps:

1) Extract PPG signal and reference BPs (SBP & DBP). Only the acceptable quality of 5s data segments was saved.
2) Pre-process PPG signal segments, including baseline removal and PPG pulse waveform normalization.
3) Extract waveform features from pre-processed PPG signal segments.
4) Train and test with 10-fold cross-validation of three different machine learning algorithms to compare the overall measurement accuracy.
5) Evaluate measurement accuracy of the three machine learning algorithms specifically for each BP category.

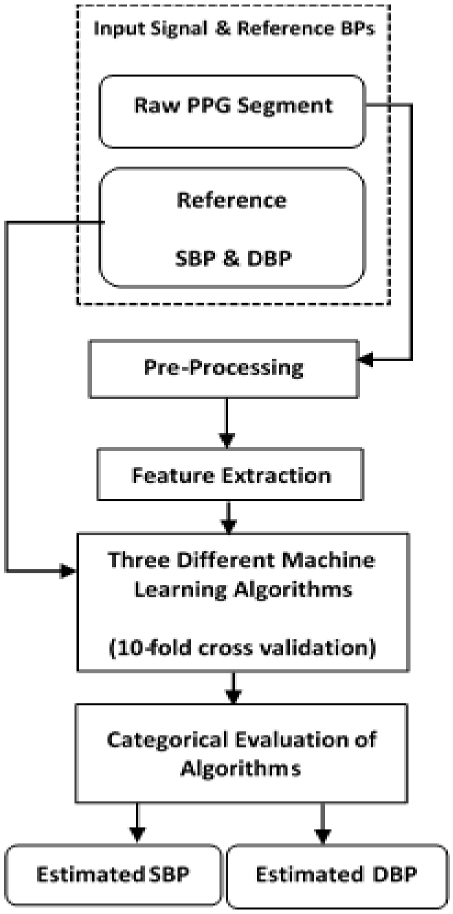
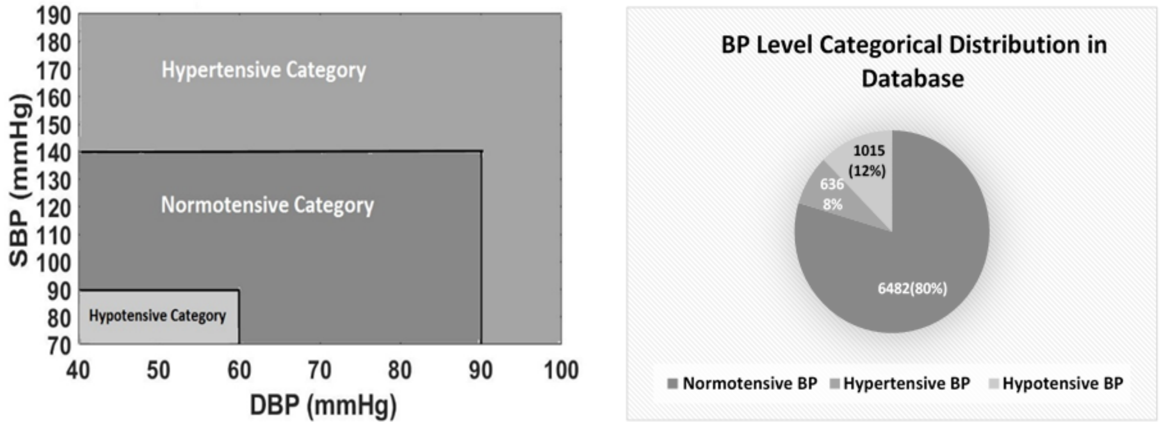



Figure 1: Flow diagram of research methodology.

**Online Database**

The University of Queensland vital signs dataset (accessed on February 2017) was used in this study. The dataset was recorded from 32 cases in Royal Adelaide Hospital using PhillipsintelliVue MP70 and MP30 with the sampling rate of 100 Hz. The signal length from each case ranged from 13 minutes to 5 hours. Raw PPG signal waveforms with their corresponding non-invasive BP (NIBP) data were extracted (Liu, Gorges et al. 2012). The length of each extracted segment was 5 second. During data segmentation, a manual check was performed to avoid unacceptable quality of PPG signal with the movement artefact and to exclude the segments without corresponding reference SBP and DBP data. The manual check was performed to ensure our machine learning models being developed did not have any interference of bad signals, allowing the BP results from different machine learning approaches to be more comparable. The number of unacceptable signal segments and the segments without reference SBP& DBP data were 9772 and 5572. Figure 3 illustrates some examples of bad quality PPG segments.

In total, as given in Table 1, 8133 signal segments of both good quality PPG and reference NIBP data were collected from the online database of 23617 signal segments. Next, each of the good quality segments was grouped into three different BP categories according to their reference BPs and the BP classification chart, as shown in Figure 2(a). The normotensive category included 6482 segments which were about 80% of the total good quality segments. The remaining Hypertensive and Hypotensive categories contained 1015 (12%) and 636 (8%), respectively as shown in Figure 2(b). Since the BPs varied during the long period of recording, each case included variable BP segments under different BP categories, as shown in Table 1.




(a) (b)

Figure 2: (a) BP classification chart to define the three BP categories and (b) categorical distribution of reference BPs of good quality PPGs in the database.

Table 1: The number of segments for each BP categorical groups, separately for each case.

| Cases | Normotensive | Hypertensive | Hypotensive | Bad quality signals | Without Reference BPs | Total |
|---|---|---|---|---|---|---|
| Case 1 | 581 | - | 64 | 419 | 376 | 1440 |
| Case 2 | 12 | - | 14 | 166 | - | 192 |
| Case 3 | 1099 | - | 50 | 1400 | 1051 | 3600 |
| Case 4 | 496 | - | - | 649 | 295 | 1440 |
| Case 5 | 50 | 32 | - | 352 | 1726 | 2160 |
| Case 6 | 357 | 56 | 69 | 158 | 80 | 720 |
| Case 7 | 128 | - | 56 | 289 | 247 | 720 |
| Case 8 | 248 | - | - | 296 | 152 | 696 |
| Case 9 | 465 | - | - | 178 | 77 | 720 |
| Case 10 | 44 | 1 | - | 195 | - | 240 |
| Case 11 | 395 | 51 | 71 | 203 | - | 720 |
| Case 12 | 312 | - | 392 | 468 | 268 | 1440 |
| Case 13 | 324 | - | 8 | 223 | 165 | 720 |
| Case 14 | - | - | 61 | 95 | - | 156 |
| Case 15 | - | - | 12 | 158 | - | 170 |
| Case 16 | - | - | 46 | 86 | 28 | 160 |
| Case 17 | 65 | - | 4 | 70 | 27 | 166 |
| Case 18 | 81 | - | - | 81 | - | 162 |
| Case 19 | 40 | 15 | - | 84 | 20 | 159 |
| Case 20 | 286 | 3 | 11 | 256 | 164 | 720 |
| Case 21 | 101 | 0 | 59 | 119 | 81 | 360 |
| Case 22 | 56 | 52 | 11 | 167 | 74 | 360 |
| Case 23 | 22 | 76 | - | 226 | - | 324 |
| Case 24 | 20 | - | - | 160 | - | 180 |
| Case 25 | 101 | 57 | 19 | 468 | 75 | 720 |
| Case 26 | 152 | 21 | 7 | 367 | 173 | 720 |
| Case 27 | 98 | - | - | 388 | 101 | 720 |
| Case 28 | 231 | - | - | 403 | 79 | 720 |
| Case 29 | 211 | - | 27 | 480 | - | 720 |
| Case 30 | 48 | - | - | 84 | - | 132 |
| Case 31 | 315 | 72 | 34 | 798 | 221 | 1440 |
| Case 32 | 144 | 200 | - | 286 | 90 | 720 |
| Total | 6482 | 636 | 1015 | 9772 | 5570 | 23617 |

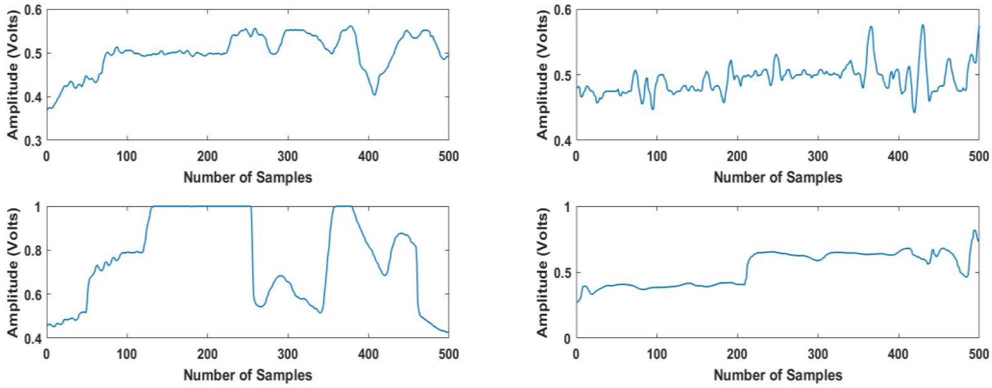



Figure 3: Examples of bad quality the PPG signal segments that cannot be processed used to extract their waveform features.

## PPG Signal Pre-Processing

Each PPG segment was firstly processed with a 4th order and 19 frame length Savitzky-Golay filter. This filter is a moving average filter to smooth PPG signal. It was selected due to the advantage of sharp edge preservation (Schafer 2011). Baseline wandering caused by the respiratory activity was also removed from the segments. The 2-Dimensional normalization (in both width & amplitude) was then performed. Figure 4 shows how a raw PPG segment is transformed to a normalised pulse. Since the reference NIBP was constant during the 5 second period of the segment, no further pre-processing of reference NIBP was required.

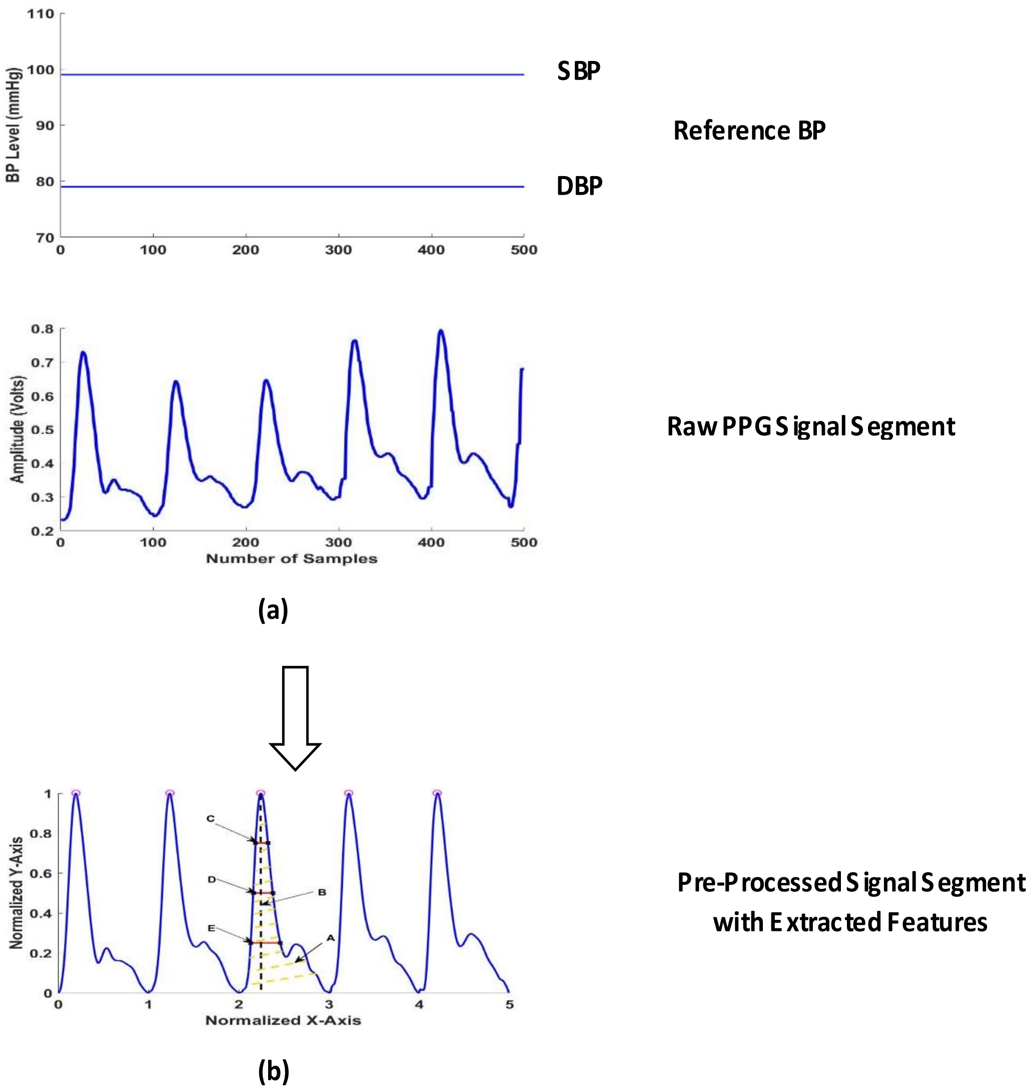



Figure 4: Illustration of pre-processing of raw PPG signals to normalised pulses, and demonstration of extracted waveform features. (a) The two horizontal straight lines are for the reference BPs and the middle sub-figure shows a 5s raw PPG signal segment. (b) Pre-processed signal segment with extracted features indicated by alphabets (*A=Pulse Area, B=Pulse Rising Time, C=Width_75%, D=Width_50%, E=Width_25%).*

## Features Extraction and Selection

Five different waveform features were initially extracted from each of the pre-processed PPG segments, which consisted of Pulse Area, Pulse Rising Time, Width 25%, Width 50% and Width 75%. The 'Pulse Area' feature of PPG segment reflects the vascular tone changes (Seitsonen, Korhonen et al. 2005). Pulse Rising Time is associated with BP changes. It has been reported that it appeared earlier in younger than in older individuals (S. R. Alty, N. Angarita-Jaimes et al. 2007). Sinha et al included this important feature in their algorithm to estimate cuffless BP (Visvanathan, Sinha et al. 2013). The PPG pulse widths are associated with the systemic vascular resistance (Awad, Haddadin et al. 2007).

To select the most significant features, the multicollinearity test was applied in this study. The presence of multicollinearity among the predictor variables affects the generalizability of the algorithm, causing a high estimated mean square error of the algorithm. An important diagnostic tool for multicollinearity among predictors, variance inflation factor (VIF) was used to determine the presence of collinearity among predictors (Schroeder, M, 1990). If VIF of a predictor larger than 10, it indicates that the predictor is highly collinear with another predictor. The most significant features were identified with multicollinearity test on the basis of their VIF. After the multicollinearity, Width_50% and Width_75% were eliminated from the training dataset due to their VIF>10.

## Machine Learning Algorithms to Estimate BPs

The training and testing dataset consisted of three most significant PPG waveform features (Pulse Area, Pulse Rising Time and Width_25%) from each of the 8133 PPG segments and their corresponding reference BPs (SBP & DBP). Due to the continuous nature of data, three commonly used regression-based machine learning algorithms were applied in this study as follows:

### Multiple Linear Regression (MLR)

MLR is a type of machine learning algorithm that has been widely used by previous researchers to estimate cuffless BP (Buxi, Redoute et al. 2015, Wang, Jia et al. 2014, Gesche, Grosskurth et al. 2012). The algorithm started with the random selection of coefficients of the linear algorithm ($\Theta_0$, $\Theta_1$, $\Theta_2$, $\Theta_3$). Each predictor was associated with a coefficient as shown in a virtual box in Figure 5a. After each iteration, the coefficients and random error ($\epsilon$, the difference between the estimated and reference BP) were updated. The least square algorithm was used to minimize the squared error as shown in equation (1). Iterative minimization of squared error continued until it converged when BP estimation was generated.

$$J\,(\Theta_0, \Theta_1, \Theta_2, \Theta_3) = \frac{1}{m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)}) \qquad (1)$$

- $h_\Theta(x) = (\Theta_0 + \Theta_1\,(Area) + \Theta_2(Crest_{Time}) + \Theta_3(\,Width\_25)) + \epsilon$

m = Total number of training data (90% of 8133)
$\epsilon$= Random error
$\Theta_{0\text{-}3}$= Coefficients
h(x) = Estimated BP
y = Reference BP

### Support Vector Machine (SVM)

SVM is a non-parametric algorithm that uses kernel function. SVM regression has a similar goal as in the least square method of MLR to minimize the error function (squared error between the estimated and reference BP). However, its approach for minimizing the function is different with MLR as it uses epsilon ($\mathcal{E}$) and the goal is to find a function whose error was no greater than $\varepsilon$. In this study, linear epsilon SVM ($\mathcal{E}$-SVM) regression which is also called L1 loss was implemented. $\mathcal{E}$-SVM has two boundaries across the hyperplane (Regression line), as shown in the line across hyperplane in Figure 5b. However, in reality, not all residuals were laid in epsilon boundary. Therefore, slack variables (another boundary) were introduced to cover all the remaining residuals, as shown in a dashed line across hyperplane in Figure 5b. Slack variables were added to make a dual objective. Each iteration updated the vectors existed in a dual objective, and the equation was analytically solved by Lagrangian function.

In SVM, the convergence criteria were based on equation (2).

$$\Delta = \frac{J(\beta) + L(\alpha)}{J(\beta)+1} \qquad (1)$$

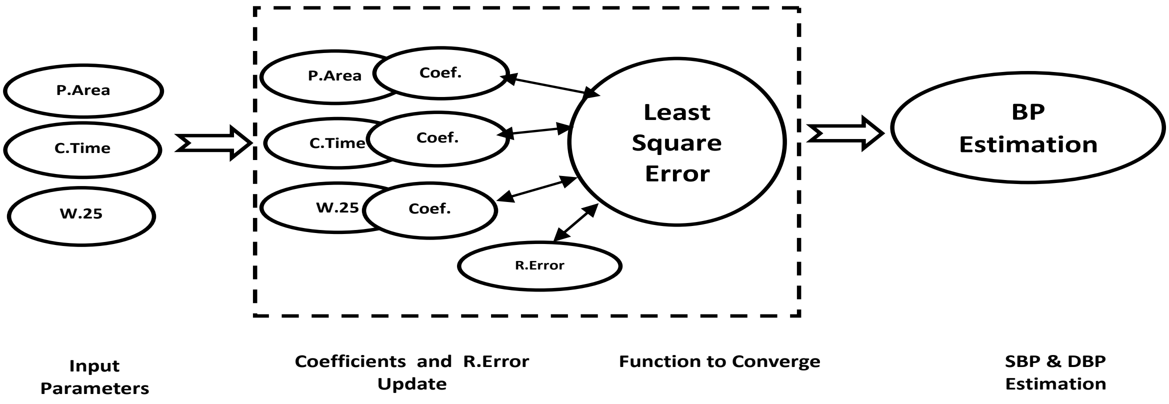
where J(β) is similar to equation (1) which in this case called primal objective. L(α) is a dual objective that was solved by the Lagrangian function. The goal was to minimize the Lagrangian function to get BP estimations. Δ represents the feasibility gap, if the feasibility gap value was less than the gap tolerance value that means algorithm considered the algorithm converged (Chen, Fan et al. 2006).

**Regression Tree**

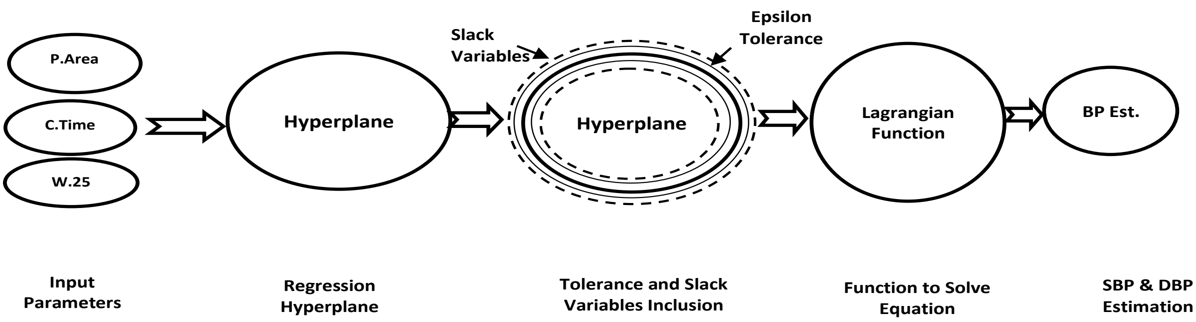
Regression Tree algorithm is another non-parametric machine learning approach for making predictions. It is a relatively fast algorithm to train the data as compared to SVM algorithm. It carries decisions from the root nodes to the leaf nodes. Regression Trees are the binary trees and the leaf that contain responses is in numeric form (Breiman 1984). It splits the data with the best optimization criteria (that subject to tree depth (α), minimum leaf size (β)) on each predictor (Pulse Area, Pulse Rising Time and Width_25%). Criteria for stopping the split to make a pure node based on the mean square error (MSE) as shown in equation (3).

*MSE (observed response) < MSE (observed response from all data) x tolerance* (2)

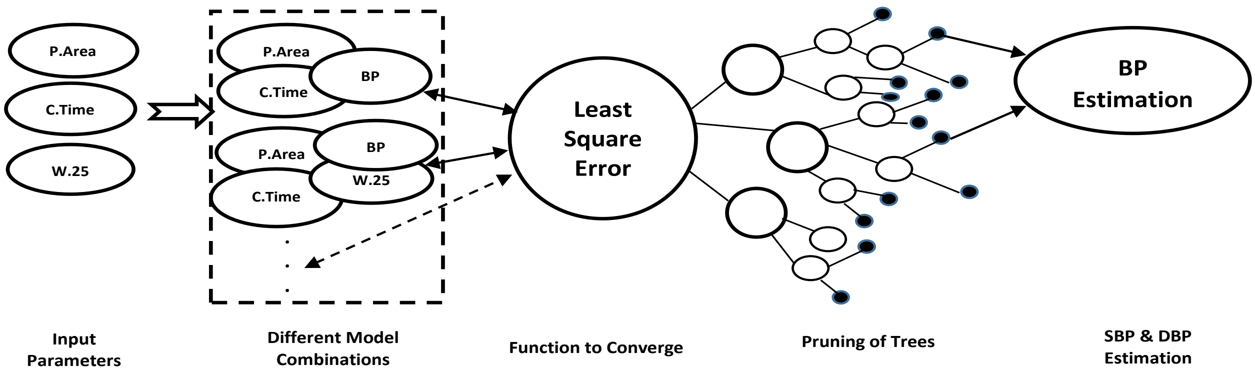
A pure node indicates that the MSE of observed response is less than the MSE of the observed response from all the data multiplied by the tolerance (Breiman 1984). For optimization, algorithm splits the branches of trees to minimize the prediction error as shown in Figure 5(c).




(a)




(b)




(c)

Figure 5:(a) Simplified flow diagrams of MLR in which coefficients and random error were updated in each iteration to converge least square error function,(b) Flow diagrams of SVM regression. Epsilon and slack variables surrounded the hyperplane were contributed to make dual objective formula with Lagrangian function to solve the equation for BP estimators and (c) Simplified flow diagrams of Regression Tree algorithms. Different algorithm combinations were used to derive least square function, prune and split tree in to branch nodes. Each node (small black colour filled circles) contains an estimation result.

### 10-fold Cross-Validation

In total, 8131×3 good quality PPG signal features and reference BPs were used to train and test the above three machine learning algorithms with 10-fold cross-validation. In each iteration, 9 folds were used to train an algorithm and the remaining fold was used to test that algorithm. The process continued until 10 iterations were completed. In the end, there was one estimated SBP and one DBP for each of the 8133 signal segments.

### Data Analysis to Evaluate Overall Measurement Accuracy

The three machine learning algorithms (Regression Tree, MLR, and SVM) were firstly evaluated in terms of overall BP measurement accuracy. After the 10-fold cross-validation of all available segments, each segment contained reference BPs (mmHg), estimated BPs (mmHg) and the difference (mmHg) between reference and estimated BP.

The averaged BPs (including both reference and estimated BPs) were calculated for each case based on all the available segments in that case. The final mean and SD of estimated BPs were then calculated for all 32 cases as an overall estimation for SBP and DBP, separately for the three machine learning algorithms. They were then compared with their reference BPs in each case to obtain overall measurement accuracy (mean difference and SD of difference).

### Data Analysis to Evaluate Measurement Accuracy in Each BP Category

For the categorical evaluation, the estimated BPs for each of the available PPG segments in each case were separated into three groups according to their reference BP category (Normotensive, Hypertensive and Hypotensive). For each case, the averaged BPs were then calculated from all the available segments under each category, which were used to obtain overall BPs across all the 32 cases, separately for each BP category. Finally, the mean difference and SD of difference between the estimated and reference BPs were calculated for each BP category and plotted using the Bland-Altman method.

## Results

### Comparison of Overall BP Measurement Accuracy

The overall BP measurement accuracy, as shown in Figure 6(a-b) and Table 2, showed that the Regression Tree a achieved the smallest mean difference of SBP (-0.1 mmHg) and SD of difference (6.5 mmHg) when compared with the MLR and SVM algorithms. Similarly, the Regression Tree achieved an acceptable mean difference (-0.6 mmHg) and SD of difference (5.2 mmHg) for DBP. It was also observed that only the Regression Tree method achieved overall acceptable accuracy to ISO standard for NIBP device validation with an average difference no greater than 5mmHg and SD no greater than 8 mmHg. Figure 6(c-h) shows the Bland-Altman plots between the reference and estimated BPs from the three machine learning algorithms.

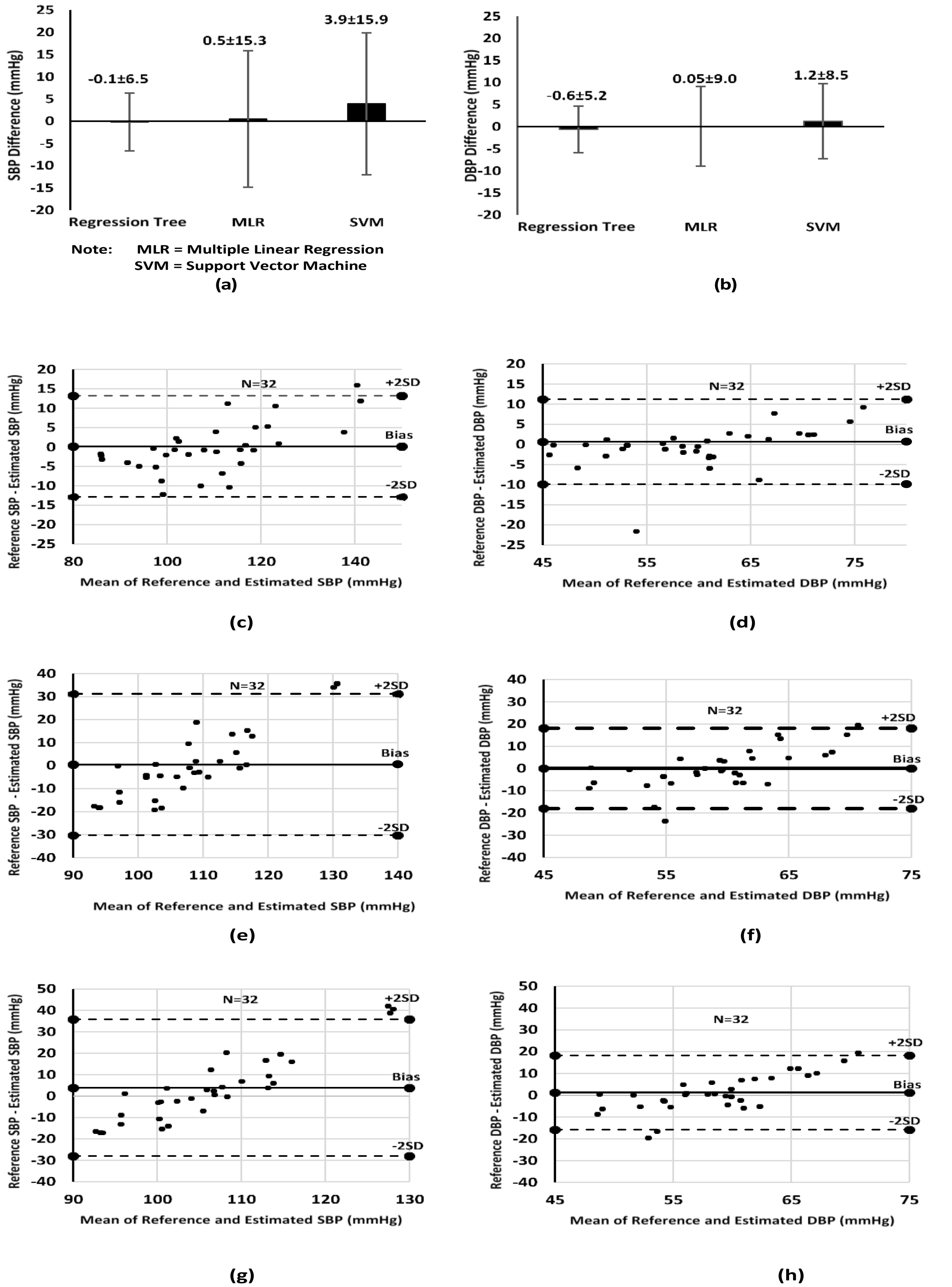



Figure 6: (a-b) Overall BP measurement accuracy from the 10-fold cross-validation, separately for the three machine learning algorithms, (c-h) Bland-Altman plots for the BPs estimated from the Regression Tree, MLR, and SVM. The left three sub-figures (c, e, g) are for SBP, and the right sub-figures (d, f, h) for DBP.

Table 2: Estimated BPs (SBP and DBP) from the Regression Tree with their corresponding reference BPs, and their difference. The results are given separately for each case.

| | SBP (mmHg) | | | DBP (mmHg) | | |
|---|---|---|---|---|---|---|
| Cases | Reference BP | Estimated BP | Difference | Reference BP | Estimated BP | Difference |
| 1 | 96.8 | 97.1 | -0.3 | 49.1 | 49.1 | 0.0 |
| 2 | 102.1 | 112.0 | -9.9 | 43.2 | 64.7 | -21.5 |
| 3 | 98.7 | 100.7 | -1.9 | 53.1 | 53.1 | 0.0 |
| 4 | 107.5 | 108.1 | -0.6 | 58.2 | 58.6 | -0.3 |
| 5 | 139.6 | 135.7 | 4.0 | 71.8 | 69.3 | 2.5 |
| 6 | 112.3 | 108.3 | 4.1 | 58.4 | 56.7 | 1.7 |
| 7 | 91.5 | 96.3 | -4.8 | 44.4 | 46.8 | -2.5 |
| 8 | 103.6 | 105.3 | -1.8 | 52.1 | 53.1 | -1.0 |
| 9 | 103.0 | 100.7 | 2.4 | 46.0 | 46.0 | -0.1 |
| 10 | 118.4 | 107.1 | 11.3 | 51.8 | 50.5 | 1.3 |
| 11 | 109.9 | 111.0 | -1.1 | 67.4 | 66.0 | 1.4 |
| 12 | 89.6 | 93.5 | -3.9 | 53.0 | 53.2 | -0.2 |
| 13 | 101.2 | 101.8 | -0.6 | 64.4 | 61.5 | 2.9 |
| 14 | 84.8 | 86.9 | -2.1 | 56.7 | 56.4 | 0.3 |
| 15 | 84.5 | 87.6 | -3.1 | 49.6 | 52.4 | -2.8 |
| 16 | 85.0 | 86.6 | -1.6 | 56.2 | 57.3 | -1.1 |
| 17 | 92.9 | 105.1 | -12.1 | 57.5 | 59.4 | -1.9 |
| 18 | 113.6 | 117.7 | -4.1 | 59.3 | 62.6 | -3.2 |
| 19 | 116.9 | 116.3 | 0.6 | 58.1 | 63.9 | -5.8 |
| 20 | 115.2 | 115.7 | -0.6 | 59.8 | 62.8 | -3.0 |
| 21 | 94.4 | 103.0 | -8.6 | 45.4 | 51.2 | -5.7 |
| 22 | 128.3 | 117.7 | 10.7 | 59.7 | 60.1 | -0.4 |
| 23 | 148.4 | 132.4 | 16.1 | 80.5 | 71.1 | 9.4 |
| 24 | 95.0 | 100.0 | -5.0 | 61.4 | 70.1 | -8.7 |
| 25 | 124.2 | 123.1 | 1.0 | 65.8 | 63.6 | 2.1 |
| 26 | 121.3 | 116.1 | 5.2 | 59.0 | 60.5 | -1.6 |
| 27 | 124.0 | 118.6 | 5.4 | 71.1 | 68.2 | 2.8 |
| 28 | 108.3 | 115.0 | -6.7 | 59.6 | 62.4 | -2.8 |
| 29 | 103.2 | 101.6 | 1.6 | 61.3 | 60.3 | 1.0 |
| 30 | 108.0 | 118.3 | -10.3 | 71.1 | 63.3 | 7.8 |
| 31 | 118.0 | 118.7 | -0.7 | 72.3 | 69.8 | 2.5 |
| 32 | 147.2 | 135.1 | 12.0 | 77.4 | 71.6 | 5.8 |
| Mean | 108.9 | 109.1 | -0.1 | 59.2 | 59.8 | -0.6 |
| SD | 16.8 | 12.8 | 6.5 | 9.4 | 7.2 | 5.2 |

**BP measurement accuracy under each BP category**

The measurement accuracies of the three machine learning algorithms under each BP category are presented in Figure 7. It can be seen that only the Regression Tree achieved acceptable accuracy to meet ISO standard for device evaluation, and it was only observed in Normotensive BP category. Its mean differences and SDs of difference for SBP and DBP were -1.1±5.7 mmHg and -0.3±5.6 mmHg. The detailed results from the Regression Tree for each BP category were presented in Table 3 and Table 4. It can be seen that Regression Tree algorithm produced higher mean differences and SD of difference under both hypertensive and hypotensive BP categories. It was also observed that, although the mean differences for the MLR and SVM algorithms were acceptable in the Normotensive category, they did not achieve acceptable ISO standard for device evaluation in terms of SD of difference, as shown in Figure 7.

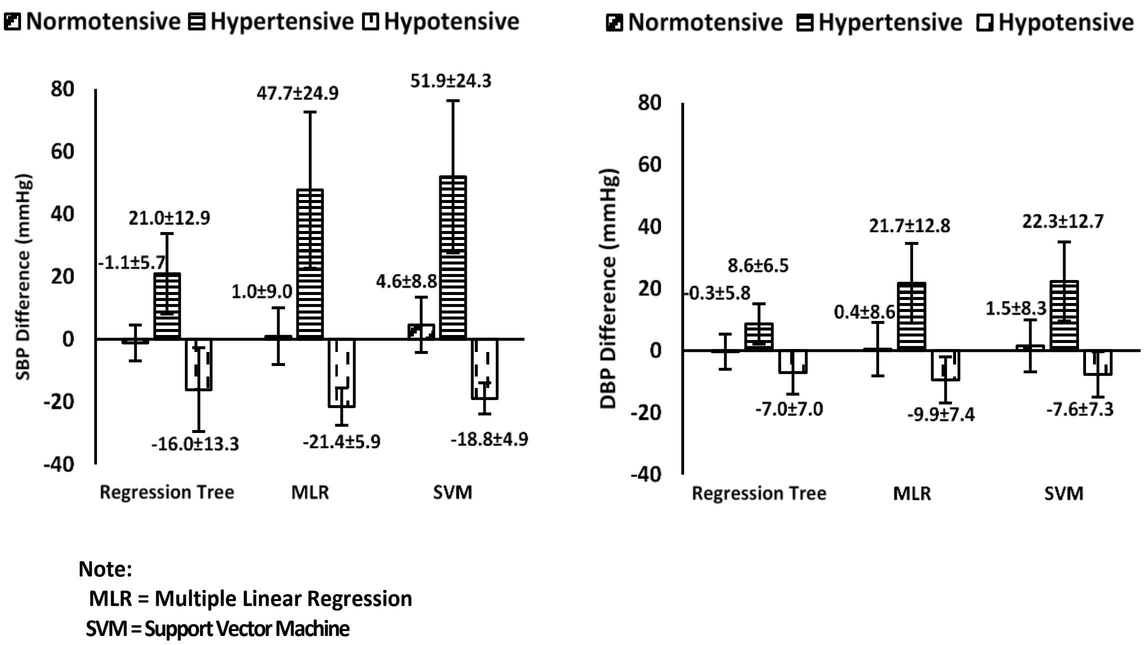



Figure 7: BP measurement accuracy under each BP category, separately for the three machine learning algorithms (Regression tree, MLR, and SVM). The data is presented with mean BP difference ± SD of BP difference.

Table 3: Estimated SBP from the Regression Tree for each individual case under the three categories, and its difference with reference SBP.

| | Normotensive | | | Hypertensive | | | Normotensive | | |
|---|---|---|---|---|---|---|---|---|---|
| Cases | Ref BP | Est BP | Diff | Ref BP | Est BP | Diff | Ref BP | Est BP | Diff |
| 1 | 98.0 | 97.6 | 0.4 | - | - | - | 85.7 | 92.4 | -6.7 |
| 2 | 120.2 | 108.2 | 12.1 | - | - | - | 85.1 | 115.7 | -30.6 |
| 3 | 99.5 | 100.7 | -1.1 | - | - | - | 80.2 | 99.1 | -18.9 |
| 4 | 107.6 | 108.2 | -0.6 | - | - | - | - | - | - |
| 5 | 121.8 | 120.4 | 1.4 | 188.7 | 185.6 | 3.2 | - | - | - |
| 6 | 108.5 | 104.5 | 4.0 | 188.9 | 155.4 | 33.5 | 84.2 | 93.2 | -9.0 |
| 7 | 97.3 | 101.6 | -4.3 | - | - | - | 87.6 | 90.7 | -3.1 |
| 8 | 104.4 | 105.9 | -1.5 | - | - | - | - | - | - |
| 9 | 105.7 | 103.2 | 2.4 | - | - | - | - | - | - |
| 10 | 117.7 | 107.0 | 10.7 | 104.0 | 104.0 | 0.0 | - | - | - |
| 11 | 111.1 | 110.2 | 1.0 | 131.9 | 130.6 | 1.3 | 87.7 | 100.5 | -12.8 |
| 12 | 103.0 | 103.3 | -0.3 | - | - | - | 78.9 | 85.7 | -6.8 |
| 13 | 101.7 | 102.2 | -0.6 | - | - | - | 84.0 | 86.5 | -2.5 |
| 14 | - | - | - | - | - | - | 84.8 | 87.0 | -2.1 |
| 15 | - | - | - | - | - | - | 84.6 | 88.2 | -3.6 |
| 16 | - | - | - | - | - | - | 85.0 | 86.6 | -1.6 |
| 17 | 93.6 | 103.9 | -10.3 | - | - | - | 82.5 | 124.4 | -41.9 |
| 18 | 113.7 | 117.1 | -3.4 | - | - | - | - | - | - |
| 19 | 106.3 | 115.9 | -9.6 | 145.0 | 117.3 | 27.8 | - | - | - |
| 20 | 116.0 | 115.8 | 0.2 | 144.0 | 105.9 | 38.1 | 87.0 | 116.0 | -29.0 |
| 21 | 99.9 | 106.3 | -6.4 | - | - | - | 85.0 | 97.8 | -12.8 |
| 22 | 116.0 | 111.6 | 4.4 | 150.4 | 124.9 | 25.5 | 87.0 | 117.4 | -30.4 |
| 23 | 123.2 | 134.1 | -11.0 | 155.7 | 131.8 | 23.9 | - | - | - |
| 24 | 95.0 | 97.7 | -2.7 | - | - | - | - | - | - |
| 25 | 119.5 | 120.2 | -0.7 | 147.8 | 127.7 | 20.1 | 83.7 | 125.9 | -42.1 |
| 26 | 117.0 | 114.9 | 2.0 | 162.0 | 124.3 | 37.7 | 87.0 | 109.5 | -22.5 |
| 27 | 124.0 | 118.6 | 5.4 | - | - | - | - | - | - |
| 28 | 106.3 | 115.2 | -8.9 | 135.5 | 121.2 | 14.3 | - | - | - |
| 29 | 105.7 | 102.9 | 2.8 | - | - | - | 83.8 | 91.8 | -8.0 |
| 30 | 108.0 | 118.3 | -10.3 | - | - | - | - | - | - |
| 31 | 114.9 | 118.5 | -3.6 | 147.0 | 125.0 | 22.0 | 85.2 | 105.7 | -20.5 |
| 32 | 118.7 | 122.7 | -4.0 | 171.6 | 146.0 | 25.6 | - | - | - |
| Mean | 109.4 | 110.5 | -1.1 | 151.7 | 130.7 | 21.0 | 84.6 | 100.7 | -16.0 |
| SD | 8.9 | 8.6 | 5.7 | 22.9 | 21.5 | 12.9 | 2.3 | 13.5 | 13.3 |

Table 4: Estimated DBP from the Regression Tree for each individual case under the three BP categories and its difference with reference DBP.

| | Normotensive | | | Hypertensive | | | Hypotensive | | |
|---|---|---|---|---|---|---|---|---|---|
| Cases | Ref BP | Est BP | Diff | Ref BP | Est BP | Diff | Ref BP | Est BP | Diff |
| 1 | 49.8 | 49.3 | 0.5 | - | - | - | 45.7 | 46.8 | -1.0 |
| 2 | 43.0 | 64.3 | -21.3 | - | - | - | 44.0 | 65.1 | -21.1 |
| 3 | 54.5 | 53.3 | 1.2 | - | - | - | 48.4 | 50.0 | -1.6 |
| 4 | 59.0 | 58.5 | 0.5 | - | - | - | - | - | - |
| 5 | 67.0 | 65.6 | 1.4 | 79.1 | 79.1 | 0.0 | - | - | - |
| 6 | 60.5 | 57.2 | 3.3 | 88.0 | 74.0 | 14.0 | 41.6 | 46.8 | -5.2 |
| 7 | 49.3 | 48.4 | 0.9 | - | - | - | 43.1 | 45.6 | -2.5 |
| 8 | 53.0 | 53.5 | -0.5 | - | - | - | - | - | - |
| 9 | 49.3 | 48.2 | 1.1 | - | - | - | - | - | - |
| 10 | 51.9 | 49.9 | 2.0 | 54.7 | 54.7 | 0.0 | - | - | - |
| 11 | 69.9 | 66.3 | 3.6 | 93.0 | 80.3 | 12.7 | 49.2 | 54.3 | -5.0 |
| 12 | 63.2 | 60.2 | 3.0 | - | - | - | 46.7 | 47.7 | -1.0 |
| 13 | 64.6 | 61.5 | 3.0 | - | - | - | 56.0 | 60.2 | -4.2 |
| 14 | - | - | - | - | - | - | 57.6 | 56.7 | 0.9 |
| 15 | - | - | - | - | - | - | 54.3 | 54.3 | 0.1 |
| 16 | - | - | - | - | - | - | 56.2 | 57.3 | -1.1 |
| 17 | 58.3 | 59.8 | -1.6 | - | - | - | 45.5 | 52.8 | -7.3 |
| 18 | 59.7 | 62.2 | -2.5 | - | - | - | - | - | - |
| 19 | 51.6 | 64.6 | -13.0 | 77.0 | 60.9 | 16.1 | - | - | - |
| 20 | 60.3 | 62.5 | -2.2 | 76.0 | 77.4 | -1.4 | 50.0 | 65.8 | -15.8 |
| 21 | 45.4 | 51.6 | -6.2 | - | - | - | 45.7 | 50.6 | -4.9 |
| 22 | 63.1 | 58.9 | 4.2 | 62.5 | 61.3 | 1.2 | 40.0 | 62.5 | -22.5 |
| 23 | 74.1 | 69.1 | 5.0 | 82.3 | 71.7 | 10.6 | - | - | - |
| 24 | 61.5 | 67.3 | -5.8 | - | - | - | - | - | - |
| 25 | 63.9 | 62.0 | 1.9 | 73.2 | 66.1 | 7.1 | 51.7 | 63.0 | -11.3 |
| 26 | 57.4 | 60.2 | -2.8 | 75.0 | 60.2 | 14.8 | 53.0 | 67.0 | -14.0 |
| 27 | 71.1 | 68.2 | 2.8 | - | - | - | - | - | - |
| 28 | 60.7 | 63.2 | -2.5 | 87.5 | 72.3 | 15.2 | - | - | - |
| 29 | 65.8 | 61.7 | 4.1 | - | - | - | 52.3 | 57.5 | -5.2 |
| 30 | 71.1 | 63.3 | 7.8 | - | - | - | - | - | - |
| 31 | 70.8 | 68.5 | 2.3 | 87.4 | 76.0 | 11.5 | 57.4 | 68.3 | -10.9 |
| 32 | 66.6 | 65.5 | 1.1 | 88.0 | 76.8 | 11.2 | - | - | - |
| Mean | 59.8 | 60.1 | -0.3 | 78.7 | 70.0 | 8.6 | 49.3 | 56.4 | -7.0 |
| SD | 8.2 | 6.3 | 5.6 | 10.9 | 8.4 | 6.5 | 5.4 | 7.4 | 7.0 |

## Discussion

In this study, the overall BP measurement accuracy from three supervised machine learning algorithms (Regression Tree, MLR, and SVM) were compared to determine which algorithm was better to estimate cuffless BPs using PPG signals only. To prevent the selection of an overfitted algorithm, the 10-fold cross-validation was used to test the overall measurement accuracy of the algorithms. The results showed that the Regression Tree achieved better overall accuracy in terms of mean and SD of BP difference as required by ISO (O'Brien, Pickering et al. 2002).

Researchers have attempted to develop MLR algorithm for PTT-based cuffless BP estimation (Mase, Mattei et al. 2011, Gesche, Grosskurth et al. 2012). Although the MLR algorithm in those studies achieved acceptable measurement accuracy, their research was still susceptible to the practical issues with two sensors for the measurement. Measurements from multiple wearable sensors could cause restricted movement and discomfort to the users (Mcdams, Krupaviciute et al. 2011). Another group also used MLR algorithm with tonometry for the estimation of cuffless BP and they succeeded to pass the ISO requirement (M. Park, H. Kang et al. 2007), but MLR is sensitive to the outliers as shown in Figure 6e, suggesting that MLR is probably not an ideal algorithm for BP estimation (Chernick 2011). In this study, SD of BP difference was higher than the requirement of no more than 8 mmHg, this was partially due to the presence of outliers.

The SVM algorithm has been used to estimate cuffless BP using heart sound signals, where acceptable BP measurement accuracy was achieved (Peng, Yan et al. 2015). Similarly in our study, SVM algorithm was applied to PPG signal features to estimate cuffless BP. However, SVM algorithm did not achieve acceptable accuracy with high SD of BP difference. The performance of SVM algorithm is mostly based on the selection of the kernel. Three different kernels (linear, Gaussian and Polynomial) have been widely used (Cristianini, 2002). In this study, the linear kernel was used to get estimation output because the selected signal features and their corresponding BPs were in linear relationships. Zhang and Feng used the same database (University of Queensland) but with different PPG signal features. They tested three machine learning algorithms MLR, Neural Network and SVM, achieved best measurement accuracy with SVM for SBP (11.6±8.2mmHg) & DBP (7.6±6.7mmHg), which were not up to the ISO standard (Zhang Y and Feng Z, 2017). Therefore, there is a need to better understand the potential reasons to improve the algorithm development.

Regression Tree algorithm is robust to the noisy data and able to make a better-fitted algorithm for discrete target data (Breiman 1984). A group of researchers used Regression Tree algorithm for PTT-based cuffless BP estimation and achieved acceptable results (Zhang, Wei et al. 2018). In this study, Regression Tree algorithm was among the best algorithm for BP estimation. The possible reason behind the success of Regression Tree is their non-vulnerability to the outliers. Another strong characteristic of this algorithm is that it also produces a well fitted algorithm in the presence of slight non-linearity within the data (Breiman 1984).

Most importantly, this study further analysed the measurement accuracy of the three machine learning algorithms under different BP categories (Normotensive, Hypertensive and Hypotensive) and found that most of the algorithms exhibited better accuracy in the Normotensive category. Previous research only presented overall BP accuracies (overall mean of difference ± SD of difference) rather than individual categorical BP accuracies (Buxi, Redoute et al. 2015, Miao, F., 2017, Wan-Hua, Hui et al. 2017). Some studies only included normotensive subjects (Wong, Pickwell-MacPherson et al. 2009, Atomi, Kawanaka et al. 2017, Shin, Min 2017). In our study, Regression Tree was found with higher mean BP difference and SD of difference in Hypertensive and Hypotensive categories in comparison with the Normotensive group. This could be caused by the low amount of data within the Hypertensive and Hypotensive categories of the online database. To make an accurate algorithm for each BP category, it is therefore suggested that specific algorithm approach for different BP categories should be considered in a future study.

This study has some limitations. Firstly, manual check to determine the quality of PPG signal segments is not practical in real scenario. The development of advanced pre-processing algorithms to automatically determine signal quality is important. It is also worth investigating the effect of noise on the estimation accuracy of machine learning models. Secondly, the training and test of the three machine learning algorithms were limited to the database of the University of Queensland. It would be useful to test the algorithms in a new database. Thirdly, due to the lack of the basic clinical variables (e.g. BMI, gender, weight, and Height) in the dataset, these variables were not included to train machine learning algorithms, which may improve the measurement accuracy of some of the algorithms (M. Park, H. Kang et al. 2007). Finally, the BP estimation was performed on the basis of each segment and only non-invasive intermittent BPs were available to be used as reference BPs to train the algorithms. In a future study, using continuous BP as reference BPs may improve the algorithms, allowing beat-to-beat BP estimation.

## Conclusions

This study developed and compared three machine learning algorithms to estimate BPs using PPG only, and revealed that the Regression Tree algorithm was the best approach with overall acceptable measurement accuracy to ISO standard for device validation. Furthermore, this study demonstrated that the Regression Tree algorithm achieved acceptable measurement accuracy only in the Normotensive category, suggesting that future algorithm development for BP estimation should be more specific for different BP categories.

## Data Availability

Data is available to access via the link below:

https://outbox.eait.uq.edu.au/uqdliu3/uqvitalsignsdataset/index.html

## Conflicts of Interest

Authors declared there are no conflicts of interest regarding the publication.


## Funding Statement

Syed Ghufran Khalid was funded by a studentship from Anglia Ruskin University.

## References


A. Visvanathan, A. Sinha and A. Pal, (2013). Estimation of blood pressure levels from reflective Photoplethysmograph using smart phones, *13th IEEE International Conference on BioInformatics and BioEngineering*, 2013, pp. 1-5.

Atomi, K., Kawanaka, H., Bhuiyan, M.S. and Oguri, K. (2017). A *Computational and mathematical methods in medicine*, 2017, pp.1-10.

Awad, A.A., Haddadin, A.S., Tantawy, H., Badr, T.M., Stout, R.G., Silverman, D.G. and Shelley, K.H, (2007). The relationship between the photoplethysmographic waveform and systemic vascular resistance. *Journal of clinical monitoring and computing*,2007,**21**(6), pp. 365-372.

Breiman, L, 1984. *Classification and regression trees.* Belmont, Calif.: Wadsworth International Group.

Buxi, D., Redoute, J.M. and Yuce, M.R, (2015). Cuffless blood pressure estimation from the carotid pulse arrival time using continuous wave radar.*Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, 2015, pp. 5704-5707.

Chen, P.H., Fan, R.E. and Lin, C.J, (2006). A study on SMO-type decomposition methods for support vector machines. *IEEE Transactions on Neural Networks*,2006,**17**(4), pp. 893-908.

Chen, W., KobayashI, T., Ichikawa, S., Takeuchi, Y. and Togawa, T, (2000). Continuous estimation of systolic blood pressure using the pulse arrival time and intermittent calibration. *Medical & biological engineering & computing*,2000,**38**(5), pp. 569-574.

Chernick, M.R., 2011. *The essentials of biostatistics for physicians, nurses, and clinicians.* Hoboken, N.J: Wiley.

Cristianini, N. and Shawe-Taylor, J. (2002). *An introduction to support vector machines*. Cambridge: Cambridge University Press.

De La Sierra, A, (2017). Ambulatory blood pressure monitoring is a useful tool for all patients. *Hipertens Riesgo vascular*, 2017, **34**(1), pp. 45-49.

Ding, X., Yan, B.P., Zhang, Y.T., Liu, J., Zhao, N. and TSANG, H.K, (2017). Pulse Transit Time Based Continuous Cuffless Blood Pressure Estimation: A New Extension and A Comprehensive Evaluation. *Scientific reports*,2017,**7**(1).

Ding, X., Zhang, Y. and Tsang, H.K, (2016). Impact of heart disease and calibration interval on accuracy of pulse transit time-based blood pressure estimation. *Physiological Measurement*,2016,**37**(2), pp. 227-237.

Gesche, H., Grosskurth, D., Kuchler, G. and Patzak, A, (2012). Continuous blood pressure measurement by using the pulse transit time: comparison to a cuff-based method. *European journal of applied physiology,* 2012, **112**(1), pp. 309-315.

Höcht, C, (2013). Blood Pressure Variability: Prognostic Value and Therapeutic Implications. *ISRN Hypertension,* 2013, pp. 1-16.

Kachuee, M., Kiani, M.M., Mohammadzade, H. and Shabany, M, (2017). Cuffless Blood Pressure Estimation Algorithms for Continuous Health-Care Monitoring. *IEEE transactions on bio-medical engineering*, 2017,**64**(4), pp. 859-869.

Liu, D., Gorges, M. and Jenkins, S.A, (2012). University of Queensland vital signs dataset: development of an accessible repository of anesthesia patient monitoring data for research. *Anesthesia and Analgesia*,2012,**114**(3), pp. 584-589. Available at: https://outbox.eait.uq.edu.au/uqdliu3/uqvitalsignsdataset/index.html [Accessed 2 Feb. 2017]

M. Park, H. Kang, Y. Huh and K. C. Kim, (2007). Cuffless and Noninvasive Measurement of Systolic Blood Pressure, Diastolic Blood Pressure, Mean Arterial Pressure and Pulse Pressure using Radial Artery Tonometry Pressure Sensor with Concept of Korean Traditional Medicine, *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* 2007, pp. 3597-3600.

Mase, M., MatteI, W., CUCINO, R., FAES, L. and NOLLO, G, (2011). Feasibility of cuff-free measurement of systolic and diastolic arterial blood pressure. *Journal of electrocardiology,* 2011,**44**(2), pp. 201-207.

Matsumura, K., Rolfe, P., Toda, S. and Yamakoshi, T. (2018). Cuffless blood pressure estimation using only a smartphone. *Scientific Reports*, 8(1).

Mccarthy, B.M., Vaughan, C.J., O'flynn, B., Mathewson, A. and O Mathuna, C, (2013). An examination of calibration intervals required for accurately tracking blood pressure using pulse transit time algorithms. *Journal of human hypertension,* 2013, **27**(12), pp. 744-750.

Mcdams, E., Krupaviciute, A., Gehin, C., Grenier, E., Massot, B., Dittmar, A., Rubel, P., Fayn, J., Wearable sensor systems: The challenges, (2011). *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011, pp. 3648-3651.

Miao, F., Fu, N., Zhang, Y., Ding, X., Hong, X., He, Q. and Li, Y. (2017). A Novel Continuous Blood Pressure Estimation Approach Based on Data Mining Techniques. *IEEE Journal of Biomedical and Health Informatics*, 2017, **21**(6), pp.1730-1740.

Nabeel, P.M., Joseph, J. and Sivaprakasam, M, (2014). Magnetic plethysmograph transducers for local blood pulse wave velocity measurement. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.IEEE Engineering in Medicine and Biology Society.Annual Conference,* 2014, pp. 1953-1956.

O'brien, E., Pickering, T., Asmar, R., Myers, M., Parati, G., Staessen, J., Mengden, T., Imai, Y., Waeber, B., Palatini, P., Gerin, W. and Working Group on Blood Pressure Monitoring of the European Society of Hypertension, (2002). Working Group on Blood Pressure

Monitoring of the European Society of Hypertension International Protocol for validation of blood pressure measuring devices in adults. *Blood pressure monitoring*, 2002, **7**(1), pp. 3-17.

Park, M., Kang, H., Huh, Y. and Kim, K.C, (2007).Cuffless and Noninvasive Measurement of Systolic Blood Pressure, Diastolic Blood Pressure, Mean Arterial Pressure and Pulse Pressure using Radial Artery Tonometry Pressure Sensor with Concept of Korean Traditional Medicine, 2007, pp. 3597-3600.

Peng, R., Yan, W., Zhang, N., Lin, W., Zhou, X. and Zhang, Y, (2015). Cuffless and Continuous Blood Pressure Estimation from the Heart Sound Signals. *Sensors (Basel, Switzerland)*,2015,**15**(9), pp. 23653-23666.

S. R. Alty, N. Angarita-Jaimes, S. C. Millasseau and P. J. Chowienczyk, (2007). Predicting Arterial Stiffness From the Digital Volume Pulse Waveform. *IEEE Transactions on Biomedical Engineering*,2007,**54**(12), pp. 2268-2275.

Schroeder, M., Lander, J. and Levine-Silverman, S. (1990). Diagnosing and Dealing with Multicollinearity. *Western Journal of Nursing Research*, 12(2), pp.175-187.

Schafer, R.W, 2011. What Is a Savitzky-Golay Filter? [Lecture Notes]. *IEEE Signal Processing Magazine*,2011,**28**(4), pp. 111-117.

Seitsonen, E.R., Korhonen, I.K., Van Gils, M.J., Huiku, M., Lotjonen, J.M., Korttila, K.T. and Yli-Hankala, A.M, (2005). EEG spectral entropy, heart rate, photoplethysmography and motor responses to skin incision during sevoflurane anaesthesia. *Acta Anaesthesiologica Scandinavica.* 2005,**49**(3), pp. 284-292.

Shin, H. and Min, S.D, (2017). Feasibility study for the non-invasive blood pressure estimation based on ppg morphology: normotensive subject study. *Biomedical engineering online*, 2017, **16**(1), pp. 10.

Shukla, S.N., Kakwani, K., Patra, A., Lahkar, B.K., Gupta, V.K., Jayakrishna, A., Vashisht, P. and Sreekanth, I, (2015). Noninvasive cuffless blood pressure measurement by vascular transit time, 2015, pp. 535-540.

Wang, R., Jia, W., Mao, Z., Sclabassi, R.J. and Sun, M, (2014). Cuff-Free Blood Pressure Estimation Using Pulse Transit Time and Heart Rate. *International conference on signal processing proceedings.International Conference on Signal Processing,* 2014, pp. 115-118.

Wan-Hua, L., Hui, W., Samuel, O.W. and Guanglin, L, (2017). Using a new PPG indicator to increase the accuracy of PTT-based continuous cuffless blood pressure estimation. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.IEEE Engineering in Medicine and Biology Society.Annual Conference,* 2017, pp. 738-741.

Watanabe, N., Bando, Y.K., Kawachi, T., Yamakita, H., Futatsuyama, K., Honda, Y., Yasui, H., Nishimura, K., Kamihara, T., Okumura, T., Ishii, H., Kondo, T. and Murohara, T, (2017). Development and Validation of a Novel Cuff-Less Blood Pressure Monitoring Device. *BTS*,2017,**2**(6), pp. 631.

Wong, M.Y., Pickwell-Macpherson, E. and Zhang, Y.T., 2009. The acute effects of running on blood pressure estimation using pulse transit time in normotensive subjects. *European journal of applied physiology,* 2009, **107**(2), pp. 169-175.

Zhang, Y., and Zhimeng, F, (2017). A SVM Method for Continuous Blood Pressure Estimation from a PPG Signal.*9th International Conference on Machine Learning and Computing*. 2017, pp. 128-132.

Zhang, B., Wei, Z., Ren, J., Cheng, Y. and Zheng, Z, (2018). An Empirical Study on Predicting Blood Pressure Using Classification and Regression Trees. *IEEE Access*,2018,**6**, pp. 21758-21768.