Limitations of Low-Cost PPG Sensors for Cuffless Blood Pressure Estimation Using IoT and Machine Learning

Ridwan Sharif [1*] and Arif Mahmud [1]

[1] Daffodil International University, Dhaka-1216, Bangladesh

{ridwan15-4381, arif.cse}@diu.edu.bd*

Abstract. One of the health issues that has attracted much attention in the world is high blood pressure (BP), which requires consistent blood pressure measurements. The traditional cuff devices, however dependable, cannot be used regularly and are inconvenient; as a result, Photoplethysmography (PPG) cuffless BP has become a popular topic in studies on Internet-of-Things (IoT) health systems. This paper gives a design and analysis of an IoT BP estimation pipeline with the use of the MAX30102 PPG sensor and an ESP32 microcontroller. Raw PPG signals were processed to eliminate noise, and then morphological, spectral, entropy, and nonlinear dynamics features were extracted. Other demographic features were also added to increase the accuracy. Mutual information, recursive feature elimination (RFE), and SHAP analysis were used to select the features. Several machine-learning models were trained and tested, such as Random Forest, CatBoost,. Through extensive feature-engineering, the highest results found were MAE = 9.08 and $R^2 = 0.24$ with Random Forest (Tuned). Due to the low sampling rate (5.6 Hz), the signal quality was poor, producing noisy and unreliable features, which could not reliably estimate the BP. This report thus indicates that cheap PPG sensors are inappropriate for clinically reliable cuffless BP prediction.

Keywords: Photoplethysmography, Blood Pressure Estimation, IoT, Machine Learning, Feature Extraction, Sensor Limitations.

1 Introduction

1.1 Introduction

Hypertension is a severe health issue of the population, which frequently causes cardiovascular morbidity and mortality globally [1][2]. A conventional cuff-based sphygmomanometer is the gold standard of non-invasive blood pressure (BP) measurement, but remains intermittent and uncomfortable, making it less efficient in continuous monitoring [3]. As a result, increased attention is paid to cuffless BP monitoring based on photoplethysmography (PPG), which is an instrument to measure volumetric blood fluctuations in the microvascular bed of tissue, using Internet-of-Things (IoT) devices and wearable health systems [4].

M. S. Arefin et al. (eds.), *Proceedings of the International Conference on Intelligent Data Analysis and Applications (IDAA 2025)*, Advances in Intelligent Systems Research 206,
https://doi.org/10.2991/978-94-6239-664-7_15

Recent research findings have proved that when used alongside sophisticated machine learning (ML) algorithms, the PPG signals can be used to predict diastolic and systolic BP with sufficient accuracy. Fiducials and morphological feature-based methods have been studied, and more recent papers have been investigating deep learning-based methods, including CNN, LSTM, and hybrid models based on attention [5][6]. Other works have merged demographic and physiological data with PPG characteristics and have demonstrated better prediction accuracy [7][8]. Most of the previous works have shown promising results, but due to the high-quality datasets or controlled laboratory conditions, they do not have a high generalizability in real-world scenarios.

The paper explores the application of the MAX30102, a ubiquitous, low-cost PPG sensor, to an IoT-based data acquisition system using the ESP32 microcontroller. An end-to-end feature extraction pipeline has been designed that consists of morphological, spectral, entropy-based features, nonlinear dynamics features, as well as demographic features (age, gender, BMI). The process of feature selection was conducted based on Mutual Information (MI), Recursive Feature Elimination (RFE), and SHAP analysis, and then several ML models were trained, namely, Random Forest, Categorical Boosting (CatBoost).

In spite of a lot of preprocessing, feature engineering, and model training and testing, the predictive performance found was poor. The best model had a mean absolute error (MAE) of about 9 mmHg and a coefficient of determination (R^2) of ~0.24. This result highlights the weakness of the MAX30102 sensor, especially its low sampling rate (~5.6 Hz) and poor signal morphology, which adds noise to the sensor and decreases the quality of features extracted.

1.2 Objective

The objective and contribution of this work are:

a) A BP estimation pipeline built on an IoT provides PPG and ML.
b) Production of valuable morphological and nonlinear features of raw PPG signals.
c) Comparison between several ML models of BP prediction.
d) Assessment of the dataset and sensor constraints as one of the important factors affecting performance.

Through this paper, we show that even though the methodology framework is sound, the limitations of low-cost hardware sensors are a major limitation for the development of accurate cuffless BP estimates. This study gives credible evidence to future researchers to use better quality sensors.

2 Literature Review and Gap Analysis

2.1 Introduction

In recent years, Cuffless blood pressure (BP) estimation on the basis of photoplethysmography (PPG) has become an active subject of research. Initial research proved the potential of the morphological and derivative PPG characteristics to be used for a continuous BP predictor [9][10]. The more modern methods have used deep learning models, such as CNN, LSTM, and hybrid architectures have demonstrated higher predictive accuracy at controlled settings [5][6]. The advantages of using such demographic characteristics as age, gender, height, weight, and BMI in combination with physiological data have also been reported positively [7][8].

Although such improvements have been made, there are still two significant limitations. First, the vast majority of studies are based on high-quality publicly available datasets like MIMIC that are obtained in clinical or laboratory settings and may not apply to real-world wearable applications. Second, there is minimal explicit literature that examines the effects of low-cost commercial sensors, which are commonly available everywhere. These hardware constraints have a direct impact on the stability of feature extraction and model performance.

2.2 Gap analysis

To emphasize this research gap, Table 1 compares the selected previous research to this study, considering feature extraction, demographic data use, sensor limitations handling, and the strategy of evaluation. As indicated, feature extraction and model comparisons are usual, but the explicit study of sensor quality as a performance determinant is peculiar to this study.

Table 1. Gap Analysis for related studies

Features	Montesinos et al. [3]	Dai et al. [5]	Mohammadi et al. [6]	Chowdhury et al. [7]	Argüello-Prada & Castaño [11]	Proposed System
Calibration-free operation	No	No	No	No	No	Yes
Standardized validation framework	No (review highlights lack)	No	No	No	No	Yes
Robustness to motion artifacts	No	No	No (trained on ICU-style data)	No	No	Yes

Features	Montesinos et al. [3]	Dai et al. [5]	Mohammadi et al. [6]	Chowdhury et al. [7]	Argüello-Prada & Castaño [11]	Proposed System
Use of demographic data	Yes (review suggests need)	No	No	Yes (age, gender used)	No	Yes (Height, Weight, Age, Gender, BMI)
Security & privacy mechanisms	No	No	No	No	No	Yes (HIPAA/GDPR)

3 Methodology

3.1 Overview

The pipeline displayed in the methodology is organized sequentially as data collection using the IoT and web interface, preprocessing, feature extraction, feature selection, model training, and evaluation. The general structure is demonstrated in Fig. 1 (Methodology Flowchart).

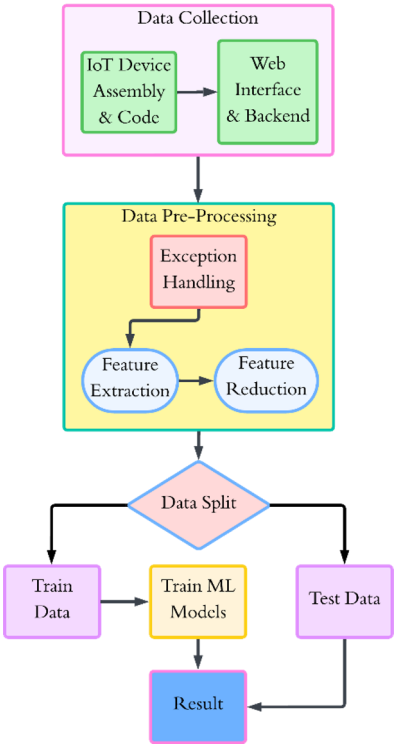

Fig. 1. Methodology Flowchart.

3.2 Data Collection & IoT Device

For data collection, a custom IoT device and web interface were developed. The IoT device consisted of a MAX30102 sensor (PPG sensor) with an ESP32 microcontroller to configure it and control the data flow. The web interface was used to collect demographic characteristics like age, gender, height, and weight, as well as to monitor and log real-time PPG data. An Omron sphygmomanometer, Omron Automatic Blood Pressure Monitor HEM-7120, was used to determine the ground truth and collect reference systolic and diastolic blood pressure values.

The use of a low-cost PPG sensor, MAX30102, revealed a great hardware limitation regarding the fidelity and sampling rate. For accurate cuffless BP prediction, research suggests a minimum sampling rate of 50 Hz and for best accuracy, a sampling rate of 100 Hz [4]. A high sampling rate is critical to ensure adequate temporal resolution to accurately capture key morphological features such as the systolic peak, dicrotic notch, and diastolic decay in the PPG waveform. Without detecting fine details, the preprocessing and feature extraction methods extract more noise than valuable data, which brings diminishing returns from machine learning models.

Despite configuring the ESP32 Microcontroller using C++ code to aim for 100 Hz, the actual stable operational rate recorded from the MAX30102 was only ~5.6 Hz. The low sampling rate progressively worsened with continuous use, resulting in the need to change the sensor multiple times during the data collection phase.

3.3 Dataset Description

An original dataset comprising 536 records was directly collected from willing volunteers in indoor environments such as university campus libraries, dormitories, and local residential spaces.

The original attributes collected consist of multimodal features:

1. Demographic Data: Age, Gender, Height, and weight were collected from volunteers, while Body Mass Index (BMI) was calculated and derived from height and weight.
2. Sensor Data: Collected using the IoT device raw Photoplethysmography (PPG) signals from the Infrared (IR). The IR PPG signal data was the primary data, which could detail the necessary key features for blood pressure prediction.

Due to how the data was collected, there was a dataset bias, primarily in:

1. Gender Imbalance: The sample exhibited a heavily skewed ratio, with almost 70% being male.
2. Age Skewness: Nearly 48% of the data was concentrated in the 20-25 age range.

This lack of demographic diversity also impacted the model learning rate and results.

3.4 Preprocessing & Feature Engineering

Since the PPG signals suffered from a low sampling rate, to circumvent this limitation and some other basic preprocessing needs, several efforts were made to:

1. Exception handling: Elimination of invalid readings (e.g., placeholders such as -999).
2. Encoding nominal data: Numerically encoding nominal data (e.g., gender).
3. Hardware replacement: multiple MAX30102 available variation was tested, but all showed similar behavior.
4. Code adjustment: The microcontroller code was adjusted, and different variations and libraries were used, but the low sample rate persisted.
5. Signal Upscaling: Digital Upscaling was considered and used to artificially increase the data resolution. However, to achieve the necessary 50 Hz or 100 Hz rate, an interpolation factor of 10 or 20, respectively. This artificially inflated sampling rate was deemed inappropriate as it injected synthetic data and ultimately did not reveal the details needed for systolic peak, dicrotic notch, and diastolic decay.
6. Feature Augmentation: The final countermeasure was to aggressively feature engineering, extracting as many attributes as possible, such as spectral, entropy-based, and morphological features, to compensate for the poor signal quality. These features were subsequently filtered using SHAP analysis and RFE.

3.5 Feature Extraction

A hybrid approach to feature extraction was adopted to represent the time-domain, frequency-domain, and nonlinear features of a signal [12].

1. Morphological Characteristics: amplitude, rise time, fall time, pulse width 25 and 75, augmentation index (AI), and dicrotic notch of timing [13].
2. Features Heart Rate Variability (HRV): mean inter-beat interval (IBI), standard deviation of NN intervals (SDNN), root mean square of successive differences (RMSSD), Poincare indices (SD1, SD2) [14].
   - Example:

$$SDNN = \sqrt{\frac{1}{(N-1)} \cdot \sum_{i=1}^{N} (NN_{\mathrm{i}} - N\bar{N})^2} \quad (1)$$

$$RMSSD = \sqrt{\frac{1}{(N-1)} \cdot \sum_{i=1}^{N-1} (NN_{i+1} - NN_i)^2} \quad (2)$$

3. Spectral Features: spectral centroid, spectral entropy, low-frequency (LF: 0.04-0.15 Hz) and high-frequency (HF: 0.15-0.4 Hz) spectral energy.
   - Spectral centroid:

$$C = \frac{\left(\sum_{k=1}^{k} f_{\mathrm{k}} \cdot P(f_{\mathrm{k}})\right)}{\left(\sum_{k=1}^{k} P(f_{\mathrm{k}})\right)} \tag{3}$$

where $P(f_k)$ Is the power at the frequency bin f_k. [4]

4. Features of Entropy: Shannon entropy and sample entropy [16].
   - Shannon entropy:

$$H = -\sum_{i=1}^{M} p(x_i) \cdot \log\left(p(x_i)\right) \tag{4}$$

$p(x_i)$ is the probability of the state x_i [15].

Nonlinear Dynamics: Hjorth parameters (activity, mobility, complexity), detrended fluctuation analysis (DFA), fractal dimensions (e.g., Petrosian FD) [16][17].

3.6 Feature Selection

Three methods of feature selection were applied to reduce redundancy and enhance generalization:

1. Mutual Information (MI): features ranked by dependency with BP values [18].
2. Recursive Feature Elimination (RFE): Pruning of weak features recursively by using a wrapper model [19].
3. SHAP values: estimated the impact of features on model predictions [20].

The last generated set contained the most informative features, and the number of features comes down to ~15-20.

3.7 Machine Learning Models

Several regression models have been trained by Linear Regression, Support Vector Regression (SVR), Random Forest, Categorical Boosting (CatBoost), Ada Boost, ANN (Artificial Neural Network), BiLSTM (Bidirectional Long-Short Term Memory), and XGBoost, K-Nearest Neighbors (KNN) Regressor.

1. CatBoost: Objective Function: gradient boosting is the minimization of the following loss:

$$L = \sum_{i=1}^{n} l\left(y_{\mathrm{i}}, F_{m-1}(x_{\mathrm{i}}) + f_{\mathrm{m}}(x_{\mathrm{i}})\right) + \Omega(f_{\mathrm{m}}) \tag{5}$$

In which F_{m-1} is the current ensemble, $f_m$is the new decision tree, l is the loss, and Ω is the regularization [21].

2. Random Forest: Random Forest constructs a set of decision trees and provides the average prediction (in the case of regression):

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T}\hat{y}_{tree_i} \tag{6}$$

T denotes the count of trees, and $h_t(x)$denotes the prediction of the t-th decision tree [22].

3. AdaBoost: AdaBoost sequences weak learners together with sample weights changing with every iteration. It predicts the last one in a weighted sum:

$$F(x) = \sum_{m=1}^{M}\alpha_{\mathrm{m}} \cdot h_{\mathrm{m}}(x) \tag{7}$$

$h_{\mathrm{m}}(x)$ represents the weak learner at iteration m, and α_{m} the weight of the weak learner.

- The weight per learner will be calculated as:

$$\alpha_{\mathrm{m}} = ½ \cdot \ln\left(\frac{(1 - \varepsilon_{\mathrm{m}})}{\varepsilon_{\mathrm{m}}}\right) \tag{8}$$

where ε_{m} is the weighted error rate of the learner h_{m} [23].

3.8 Evaluation Metrics

There were five evaluation metrics used to evaluate model performance:

1. Mean Absolute Error (MAE) [24]:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y_i}| \tag{9}$$

2. Mean Squared Error (MSE) [24]:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i})^2 \tag{10}$$

3. Root Mean Squared Error (RMSE) [24]:

$$RMSE = \sqrt{MSE} \tag{11}$$

4. Coefficient of Determination (R^2) [25]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{12}$$

5. Mean Absolute Percentage Error (MAPE) [26]:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{y_i} \times 100 \quad (13)$$

These metrics gave a detailed assessment of forecasting and model generalization.

4 Results and Analysis

4.1 Overview of the model performance

Each of the models was trained and tested on the preprocessed, feature-extracted, and feature-selected data based on the metrics mentioned in the subchapter "3.8 Evaluation Metrics", and Table 2 summarizes the overall performance of each of the models. The models, Random Forest (tuned) and CatBoost (tuned), had the best performance, considering MAE ≈ 9 mmHg and $R^2 \approx 0.24$, whereas the SVR and BiLSTM had the poorest results. Deep learning models, ANN, and BiLSTM underperformed compared to other Machine learning models due to the small size of the dataset.

Table 2. Performance Comparison of Models

Model	MAE	MSE	RMSE	R^2	MAPE (%)
Linear Regression	9.4	175.7	13.25	0.190	9.17
Random Forest (Tuned)	9.08	165.7	12.87	0.241	8.83
XGBoost	9.78	185.2	13.61	0.146	9.61
ANN	9.45	168.8	12.99	0.213	9.19
BiLSTM	10.21	198.6	14.09	0.107	10.03
KNN	9.73	181.1	13.45	0.175	9.52
SVR	9.82	194.1	13.93	0.137	9.49
AdaBoost Regression	9.84	170.2	13.04	0.230	9.87
CatBoost (Tuned)	9.20	167.6	12.94	0.241	9.02

4.2 Visualization of Best models

To gain a better understanding, scatter plots and residual plots have been drawn.

4.2.1 CatBoost (Tuned)

The true vs. predicted plot (Fig. 2) demonstrates that predictions coincide with the real BP values rather moderately, but strong deviations are observed at higher pressures.

The residual distribution (Fig. 3) is more or less symmetrical at zero, which means that there is not much bias, but large variance.

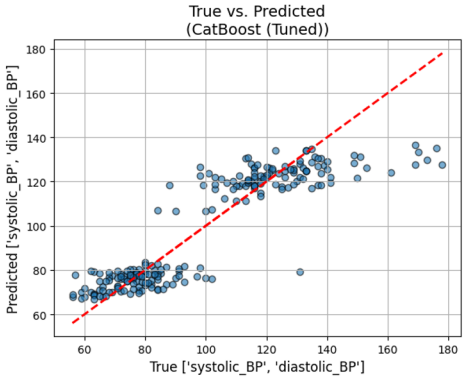

Fig. 2. True vs. Predicted (CatBoost Tuned).

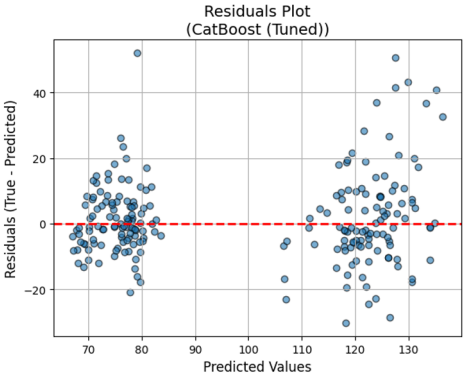

Fig. 3. CatBoost (Tuned) Residuals Plot.

4.2.2 Random Forest (Tuned)

Random Forest had a slightly smaller MAE (9.08 mmHg) and resulted in smaller scatter (Fig. 4) and residuals (Fig. 5), which implies more generalization compared to Cat-Boost. The accuracy curve (Fig. 6) shows that this error margin represents a higher percentage of predictions within +-10 mmHg of the true BP, which is better than the other models.

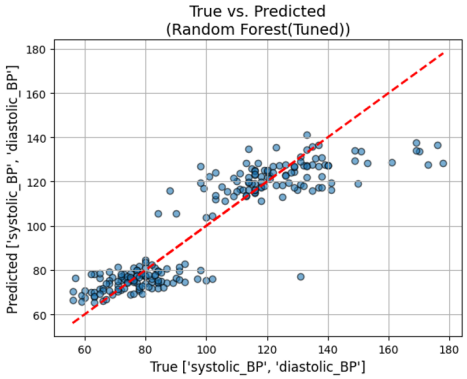

Fig. 4. True vs. Predicted (Random Forest Tuned).

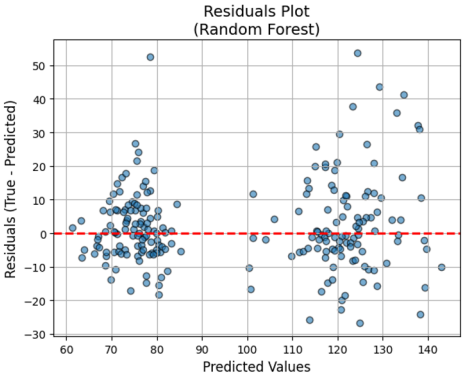

Fig. 5. Random Forest (Tuned) Residuals Plot.

4.3 Comparative Visualization Among Models

Bar graphs of MAE & MAPE and R^2 have been plotted to compare all models (Fig. 6). The charts clearly demonstrate high performance of the two models, namely Random Forest and CatBoost, with the largest errors produced by SVR and BiLSTM.

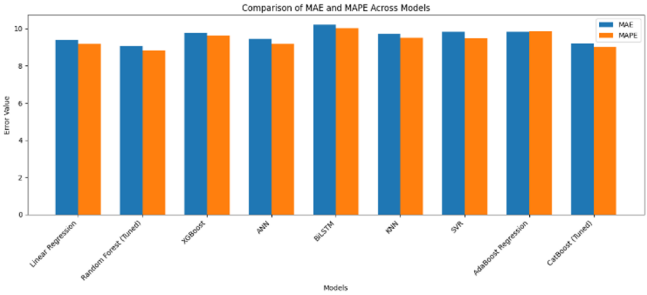

Fig. 6. (A) Performance of all the models in comparison: MAE & MAPE

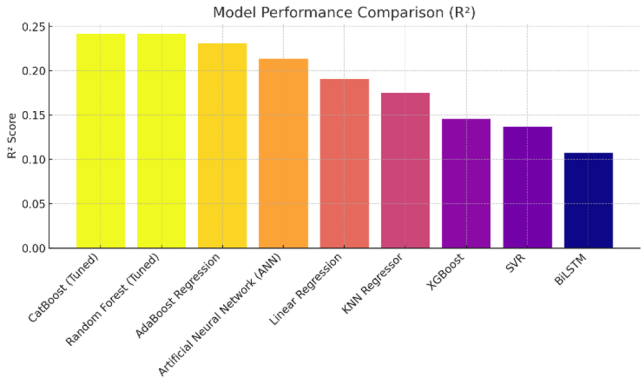

Fig. 6. (B) Performance of all models in comparison: R^2

4.4 Discussion and Observed Limitations

The comparison analysis establishes the fact that ensemble-based models, especially the Random Forest and CatBoost, performed better than the linear and shallow models to identify nonlinear relationships in the data. Nevertheless, even with much preprocessing and feature engineering, none of the models achieved any clinically acceptable performance levels (MAE $\leq$ 5 mmHg, $R^2 > 0.5$). The optimal values leveled off at MAE $\approx$ 9 mmHg and $R^2 \approx 0.24$, which is not adequate to be used in clinical practice.

This deficit is mostly since the data acquisition hardware and dataset quality are limited:

1. Limitations of Sensors: The MAX30102 sensor, being cheap and commonly used in IoT, has a low sampling rate (≈5.63 Hz) and poor morphology of the

waveforms. This complicates the accurate representation of fiducial points, including systolic peaks, dicrotic notches, and diastolic properties, resulting in unsteady morphological and spectral descriptions.

2. Sensitivity to Noise: Motion artifacts and baseline drift were major factors that affected entropy and nonlinear dynamics properties, which necessitate stable and well-defined beat-to-beat waveforms. These features, therefore, resulted in variation instead of raising the accuracy of prediction.
3. Dataset Bias: The sample was not diverse, as around 70 percent of the participants were male, and almost half of the sample fell in the age bracket of 20-25. Such a bias lowered the generalization ability of models, and it would be hard to estimate BP with underrepresented populations.
4. Poor Sample Size: The small sample size (n=536) made deep learning models like BiLSTM inaccessible to temporal dependencies, which caused the underfitting or overfitting.

Collectively, these results emphasize that although the pipeline and methodology were valid, the quality of the hardware, mainly the sensor , and the lack of diversity in the dataset were the constraining factors, rather than the pipeline or the machine learning model. This paper thus adds to the literature that low-cost sensors cannot be used to estimate clinically robust cuffless BP, with new research needing to concentrate on higher-quality sensors and more balanced data.

5 Conclusion

5.1 Summary

This paper has proposed and tested an IoT-based cuffless blood pressure (BP) estimation system using photoplethysmography (PPG) measurements of the MAX30102 sensor. An extensive feature extraction model was executed and used to incorporate morphological, spectral features, entropy-based features, and nonlinear features, and demographic variables. Mutual Information, Recursive Feature Elimination, and SHAP analysis were used to select features. Several machine learning models, such as Random Forest, CatBoost, BiLSTM, and others, were trained and tested on MAE, RMSE, R^2, and MAPE.

The most successful models were Random Forest (tuned) and CatBoost (tuned) with MAE ~9mmHg and $R^2 \approx 0.24$. Although these findings prove that the proposed pipeline is a viable project, they also indicate that low-cost sensors are not capable of delivering clinically acceptable accuracy.

5.2 Limitation

The results highlight how the bottleneck of machine learning models is not the major problem, but the quality of the sensor and the dataset:

1. The low sampling rate (≈5.63 Hz) and poor morphology of the waveforms of the MAX30102 sensor created noise, which restricted the stability of the morphological and entropy-based features.
2. The sample size (n=536) was small and non-representative, consisting of about 70% males and about 48% of young adults (20-25), which limited its generalization ability.
3. A large constraint for this research was the inaccessibility of high-quality and medical-grade PPG sensors and other reference devices within the locality (Bangladesh). This unavailability prevented cross-sensor testing.
4. The data constraints limited the nonlinear and deep learning models, which resulted in overfitting or underfitting.

5.3 Future work

To obtain a correct and clinically feasible estimation of cuffless BP, future research should combat such limitations by:

1. Using better PPG sensors with better sampling rates and better peak detectors.
2. Gathering more varied and bigger data sets of various age groups, genders, and health conditions.
3. Investigating sensor fusion strategies, including using PPG with ECG or accelerator data, to enhance resistance to noise and motion artifact.
4. The creation of non-calibration algorithms to decrease dependence on cuff-based reference measurements.
5. Future study dedicated to cross-sensor examination to find the best PPG sensor for cuffless BP prediction. The testing of various types of PPG sensors of high-quality and medical grade, as well as reevaluating cheap options.

Through these issues, future studies would be able to utilize the framework described to facilitate the creation of reliable, non-invasive, and real-time BP monitoring in IoT-based healthcare systems.

References

1. Mukkamala R, Shroff SG, Kyriakoulis KG et al (2025) Cuffless blood pressure measurement: Where do we actually stand? Hypertension 82(6):957–970.
2. Henry B, Merz M, Hoang H et al (2024) Cuffless blood pressure in clinical practice: Challenges, opportunities and current limits. Blood Pressure 33(1):2304190.
3. Montesinos L, Martinez-Rios E, Alfaro-Ponce M, Pecchia L (2021) A review of machine learning in hypertension detection and blood pressure estimation based on clinical and physiological data. Artif Intell Med 118:102125.
4. El-Hajj C, Kyriacou PA (2020) A review of machine learning techniques in photoplethysmography for the non-invasive cuff-less measurement of blood pressure. Biomed Signal Process Control 62:102132.
5. Dai D, Ji Z, Wang H (2024) Non-invasive continuous blood pressure estimation from single-channel PPG based on a temporal convolutional network integrated with an attention mechanism. Appl Sci 14(14):6061.
6. Mohammadi H, Tarvirdizadeh B, Alipour K, Ghamari M (2025) Cuff-less blood pressure monitoring via PPG signals using a hybrid CNN-BiLSTM deep learning model with attention mechanism. Sci Rep.
7. Chowdhury MH, Shuzan MNI, Chowdhury ME et al (2020) Estimating blood pressure from the photoplethysmogram signal and demographic features using machine learning techniques. Sensors 20(11):3127.
8. Wong MKF, Hei H, Lim SZ, Ng EYK (2023) Applied machine learning for blood pressure estimation using a small, real-world electrocardiogram and photoplethysmogram dataset. Math Biosci Eng 20(1):975–997.
9. Liu M, Po LM, Fu H (2017) Cuffless blood pressure estimation based on photoplethysmography signal and its second derivative. In: 2017 E-Health and Bioengineering Conference (EHB), pp 555–558.
10. Zadi AS, Alex R, Zhang R et al (2018) Arterial blood pressure feature estimation using photoplethysmography. arXiv preprint arXiv:1811.06039.
11. Argüello-Prada EJ, Castaño Mosquera CD (2025) Exploring supervised machine learning models to estimate blood pressure using non-fiducial features of the photoplethysmogram (PPG) and its derivatives. Phys Eng Sci Med 1–16.
12. Park J et al (2022) Photoplethysmogram Analysis and Applications. Front Physiol 12:808451.
13. Mehrabbeik M, Rashidi S (2019) Estimation of Cuffless Systolic and Diastolic Blood Pressure Using Pulse Transient Time. In: Proceedings of the 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), pp 1–6.
14. Tsai YY et al (2025) Photoplethysmography-based HRV analysis and machine learning for blood pressure estimation. Front Physiol 13:11970940.
15. Zhu Z et al (2025) An emotion recognition method based on frequency-domain features of PPG signals. Front Physiol 16:1486763.
16. Markov K et al (2025) Interpretable feature-based machine learning for automatic classification of stress using physiological signals. npj Digit Med 8:41.
17. Charlton PH et al (2021) Photoplethysmography Signal Processing and Synthesis. In: Photoplethysmography: Technology, Applications and Clinical Perspectives, pp 123–145.
18. Beraha M et al (2019) Feature Selection via Mutual Information: New Theoretical Insights. arXiv preprint arXiv:1907.07384.
19. Awad M et al (2023) Recursive Feature Elimination with Cross-Validation Using Decision Tree Model as Estimator. MDPI Appl Sci 12(5):67.

20. Kraev E et al (2024) Shap-Select: Lightweight Feature Selection Using SHAP Values and Regression. arXiv preprint arXiv:2410.06815.
21. Prokhorenkova LV et al (2018) CatBoost: Unbiased Boosting with Categorical Features. arXiv preprint arXiv:1706.09516.
22. Breiman L (2001) Random Forests. Mach Learn 45(1):5–32.
23. Freund Y, Schapire RE (1997) A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. J Comput Syst Sci 55(1):119–139.
24. Willmott CJ, Matsuura K (2005) Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing Average Model Performance. Clim Res 30(1):79–82.
25. Cohen J (1988) Statistical Power Analysis for the Behavioral Sciences, 2nd edn. Lawrence Erlbaum Associates.
26. Makridakis S et al (1993) Accuracy Measures: Theoretical and Practical Concerns. Int J Forecast 9(4):527–529.