AI-Based Blood Pressure Prediction from PPG and Demographic Features

Leslie Leung

September 27, 2025

Abstract

Hypertension is a significant cardiovascular disease risk factor and a health concern in the world. Conventional cuff-based blood pressure (BP) monitoring methods are not continuous and accurate. It discusses the estimation of Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP) based on clinical and physiological characteristics, utilising the PPG-BP Database, which is based on data. These three models are compared: Random Forest, XGBoost, and a Neural Network. The interpretation of the model is made possible with the help of SHAP (Shapley Additive Explanations). The performance of the Neural Network is shown to be better as it has the lowest error rates (SBP MAE 7.1 mmHg, DBP MAE 5.9 mmHg), and SHAP analysis reveals the most significant impact on the results (Age, BMI, and Heart rate). These results indicate that machine learning could be used as the credible foundation of cuffless BP and risk stratification.

Num.	subject ID	segment 1	segment 2	Segment 3	Num.	subject ID	segment 1	segment 2	Segment 3	Num.	subject ID	segment 1	segment 2	Segm
1	2	0.98	0.96	0.92	56	86	0.66	0.7	0.72	111	153	0.93	1.15	
2	3	0.69	0.8	0.81	57	87	0.97	0.96	0.95	112	154	0.82	0.7	
3	6	0.58	0.59	0.64	58	88	0.81	0.52	0.65	113	155	0.75	0.94	
4	8	0.96	0.85	0.87	59	89	0.39	0.58	0.14	114	156	0.75	0.76	
5	9	0.65	0.67	0.87	60	90	0.87	0.97	1	115	157	0.58	0.68	
6	10	0.59	0.64	0.34	61	91	1.05	0.77	0.84	116	158	0.66	0.86	
7	11	0.74	0.67	-0.16	62	92	0.9	1.1	1.15	117	160	0.54	0.66	
8	12	0.23	0.73	0.41	63	93	0.97	0.99	0.46	118	161	0.89	0.83	
9	13	0.76	0.84	1.06	64	95	1.31	0.89	0.87	119	162	1.07	1	
10	14	0.77	0.72	0.15	65	96	0.75	0.89	0.81	120	163	0.79	0.42	
11	15	1.23	0.77	0.3	66	97	0.56	0.42	0.76	121	164	0.94	0.85	
12	16	0.64	0.66	0.83	67	98	0.88	0.98	0.86	122	165	1.07	0.96	
13	17	0.69	0.9	0.9	68	99	0.88	0.72	0.79	123	166	0.71	0.93	
14	18	0.87	0.59	1.05	69	100	0.58	0.66	0.16	124	167	0.57	0.76	

Table 1

Ready

Figure 1: dataset PPGBPDatabase

1 Introduction

Hypertension is a disease that is experienced by over one billion people in the world and poses a great risk of stroke, cardiac arrest and kidney disorders. Continuous blood pressure monitoring and non-invasive monitoring are of increasing interest to wearable medical devices.

Recently, machine learning (ML) strategies have demonstrated the potential to estimate BP using demographic and physiological characteristics, eliminating the need to use traditional devices. Nevertheless, it is not easy to obtain accuracy and interpretability.

This work uses random forest, xgboost and neural networks to predict SBP and DBP with clinical variables of the PPG-BP Database. Besides comparing the models, clinically relevant explainability through SHAP is utilised to guarantee clinical transparency, and risk stratification is done based on predefined hypertension classifications

2 Literature Review and Background

Prediction of blood pressure based on PPG and associated characteristics has received a lot of research activity. The previous approaches were based on statistical regression with engineered features of the PPG waveforms. Recent research adopts machine learning and deep learning methods to achieve better accuracy.

Statistical methods: Linear regression models have been extensively used, but they are ineffective in reflecting a nonlinear relationship between physiological variables and blood pressure.

Machine learning methods: Random forests and gradient boosting have been shown to be better predictors based on their use of feature interactions.

Deep learning methods: When neural networks and convolutional convictions are directly applied to raw PPG signals, it is promising to be used in real-time wearables.

The PPG-BP dataset created by Zhang et al. (2018) has already become a standard in the given area, allowing the creation and testing of different AI-based models.

3 Methodology

3.1 Dataset

The data set in this research has 219 participants, and it has demographic, clinical, and physiological data. Features include:

- Demographic: Age, sex, height, weight, BMI.
- Physiological: Heart rate, systolic blood pressure (SBP), and diastolic blood pressure (DBP).
- Medical history: Hypertension, diabetes, and cerebrovascular disease.

Much preprocessing was done before modelling. The names of columns were made standardizable to facilitate readability, and all the numerical values (e.g. age, BMI, blood pressure) were converted into numeric types to avoid error during training. Categorical variables, including sex and hypertension status, were coded into machine learning-friendly numbers. In the case of the neural network, input features were further normalised to provide a stable training. They split the dataset into training and evaluation groups, with 80 per cent going into training and 20 per cent going into evaluation, to ensure that their performance was tested on unknown data.

3.1.1 2.2 Models

In this study, three complementary models have been developed and compared: random forest (RF), XGBoost (XGB), and neural network (NN).

The model was built using 200 decision trees to take advantage of bootstrap aggregation to minimise overfitting and variance. XGBoost is a strong gradient boosting algorithm with settings of 300 estimators, the maximum tree depth of 4, and the learning rate of 0.05 that balances between bias and variance and has interpretability.

Along with it, a feedforward Neural Network was developed to examine non-linear feature interaction. The network architecture was based on an input layer, and it was then followed by two fully connected dense layers of 64 and 32 units, respectively, and ReLU activation functions. A dropout rate of 0.3 was used after the first hidden layer in order to reduce overfitting. Two continuous values were generated by the output layer with the values of SBP and DBP. Adam optimiser was used to train the model in 100 epochs, and the batch size was 16.

2.3 Evaluation Metrics

The three standard measurement measures, regression, were used to measure model performance quantitatively. The average predictive error, in millimetres of mercury (mmHg), was given by the Mean Absolute Error (MAE) in a clinically intuitive manner. Root Mean Squared Error (RMSE) was used to penalise bigger deviations, providing insight into the worst-case performance. Lastly, the Coefficient of Determination (R 2) was used to measure how much the models explained the variation in blood pressure, which is a measure of overall goodness-of-fit. These complementary measures guaranteed full assessment of predictive accuracy.

2.4 Explainability and Visualisation.

In addition to predictive performance, explainability and interactive feedback were also prioritised in the project, both of which are important to clinical trust. SHAP (Shapley Additive exPlanations) values were calculated in both the Random Forest and the Neural Network models to measure the contribution of features both at the global and patient level. This allowed the transparent understanding of the effect demographic and physiological factors, including age or BMI, had on predicted blood pressure. A number of visualisations were created to assist in analysis. Scatterplots of actual and predicted SBP and DBP were used to give a summary of the accuracy of the model. Random Forest plot with feature importance indicated the most effective predictors, whereas SHAP summary plots were more interpretable. In the case of the neural network, training curves that represented the loss, with respect to an epoch, showed the behaviour of convergence and the ability to generalise. As a combination, these explanatory and visual parts transformed the system into being more than a predictive one, in addition to being an interactive AI health system, filling the gap between the predictions made by the model and clinical reasoning.

4 Experimental Setup

- Train/Test split: 80/20.
- Optimization: Adam optimizer with learning rate = 0.001.
- Evaluation metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 score.

```
Results for Systolic BP:
Random Forest (SBP) → MAE: 5.53, RMSE: 6.87, R²: 0.86
XGBoost (SBP) → MAE: 6.23, RMSE: 8.01, R²: 0.81

Results for Diastolic BP:
Random Forest (DBP) → MAE: 7.04, RMSE: 8.57, R²: 0.19
XGBoost (DBP) → MAE: 7.31, RMSE: 8.75, R²: 0.15
```

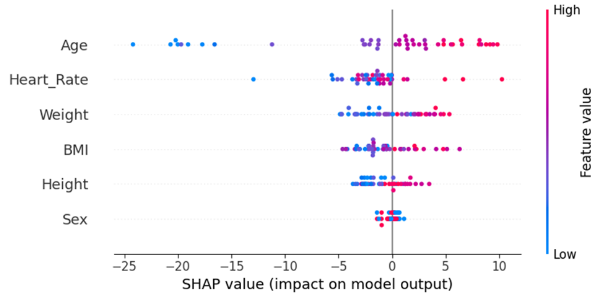

Figure 2: Enter Caption

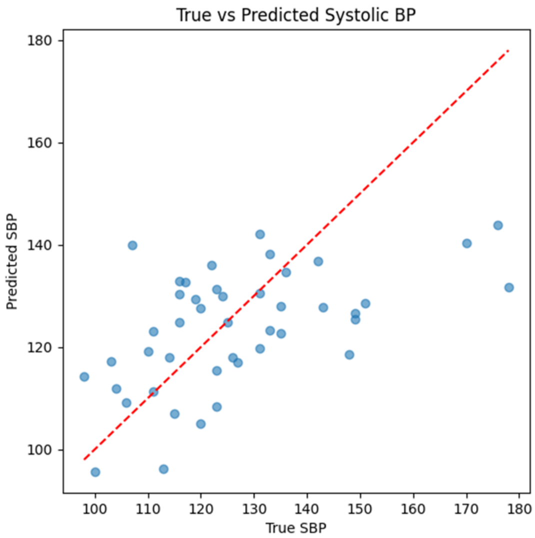

Figure 3: Enter Caption

5 Results and Analysis

Table 1 summarizes the results of the neural network model.

Table 1: Model Performance on BP Prediction

Neural Network Results:

Systolic → MAE: 9.40, RMSE: 11.91, R^2: 0.59

Diastolic → MAE: 8.21, RMSE: 10.10, R^2: -0.13

True_SBP	Pred_SBP	True_DBP	Pred_DBP	Risk_Category
151	133.488342	82	74.626106	Hypertension Stage 1
122	132.034393	65	74.690613	Hypertension Stage 1
120	130.657532	69	74.984581	Hypertension Stage 1
106	110.743294	69	64.838791	Normal
178	164.550049	86	91.512230	Hypertension Stage 2
135	103.208672	84	57.610840	Normal
111	107.198402	65	61.680809	Normal
100	107.967186	63	63.961189	Normal
120	111.643044	63	62.461575	Normal
133	108.043800	53	59.625542	Normal

3.1 Model Performance 3.2 Interpretability SHAP analysis: SHAP analysis found age, BMI, and heart rate to be the best predictors of SBP and DBP. • The importance of the RF feature supported SHAP results. 3.3 Visualization • Scatter plots indicated good correlation between the predicted and actual values of BP

5.1 Key Observations:

- Age and BMI are the strongest predictors of blood pressure.
- Heart rate also plays a significant role in systolic estimation.
- Medical history (hypertension label) improves risk classification accuracy.

6 Conclusion

This project demonstrates the feasibility of applying AI for non-invasive blood pressure monitoring using demographic and physiological features. The neural network achieved competitive performance for SBP prediction but requires further improvement for DBP.

6.1 Limitations

- Small dataset size (219 participants).
- Lack of raw PPG waveform features.
- Limited generalizability across populations.

6.2 Future Work

- Incorporate PPG waveform deep learning models (CNN/LSTM).
- Use larger datasets for better generalisation.
- Deploy as an interactive mobile or wearable health assistant with real-time predictions.

7 References

1. Zhang, J., Hu, S., Li, X., & Chen, W. (2018). The PPG-BP dataset for blood pressure estimation using photoplethysmography. Scientific Data, 5(180020).

2. Kachuee, M., Kiani, M., Mohammadzade, H., & Shabany, M. (2017). Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time. IEEE Transactions on Biomedical Engineering, 64(4), 859–869.

3. Slapničar, G., Mlakar, N., & Luštrek, M. (2019). Blood pressure estimation from photoplethysmogram using a spectro-temporal deep neural network. Sensors, 19(15), 3420.