

## Article

# Cuffless Blood Pressure Measurement Using Linear and Nonlinear Optimized Feature Selection

Mohammad Mahbubur Rahman Khan Mamun \* and Ali T. Alouani 

Department of Electrical and Computer Engineering, Tennessee Technological University,  
Cookeville, TN 38501, USA; aalouani@tntech.edu

\* Correspondence: mrkhanmamu42@tntech.edu

**Abstract:** The cuffless blood pressure (BP) measurement allows for frequent measurement without discomfort to the patient compared to the cuff inflation measurement. With the availability of a large dataset containing physiological waveforms, now it is possible to use them through different learning algorithms to produce a relationship with changes in BP. In this paper, a novel cuffless noninvasive blood pressure measurement technique has been proposed using optimized features from electrocardiogram and photoplethysmography based on multivariate symmetric uncertainty (MSU). The technique is an improvement over other contemporary methods due to the inclusion of feature optimization depending on both linear and nonlinear relationships with the change of blood pressure. MSU has been used as a selection criterion with algorithms such as the fast correlation and ReliefF algorithms followed by the penalty-based regression technique to make sure the features have maximum relevance as well as minimum redundancy. The result from the technique was compared with the performance of similar techniques using the MIMIC-II dataset. After training and testing, the root mean square error (RMSE) comes as 5.28 mmHg for systolic BP and 5.98 mmHg for diastolic BP. In addition, in terms of mean absolute error, the result improved to 4.27 mmHg for SBP and 5.01 for DBP compared to recent cuffless BP measurement techniques which have used substantially large datasets and feature optimization. According to the British Hypertension Society Standard (BHS), our proposed technique achieved at least grade B in all cumulative criteria for cuffless BP measurement.

**Keywords:** electrocardiogram; photoplethysmography; symmetric uncertainty; fast correlation; ReliefF algorithm; blood pressure measurement



**Citation:** Khan Mamun, M.M.R.; Alouani, A.T. Cuffless Blood Pressure Measurement Using Linear and Nonlinear Optimized Feature Selection. *Diagnostics* **2022**, *12*, 408. <https://doi.org/10.3390/diagnostics12020408>

Academic Editor: Nikolaos Fragakis

Received: 15 January 2022

Accepted: 30 January 2022

Published: 5 February 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Cardiovascular diseases (CVD) are responsible for major health concerns around the world. Potential symptoms or information which may provide early prediction of CVD may help many people to take proper precautionary measures [1,2]. Consistent high blood pressure (BP) is one of the four major risk factors for CVD and the main reason for hypertension, so frequent monitoring and tracking of blood pressure trends should be prioritized [3]. Chronic hypertension may cause damage to several organs [4]. So, the early detection of hypertension allows patients to take treatment at an early stage to reduce the possibility of deteriorating cardiac condition due to hypertension which does not show symptoms at an early stage. The common BP measurement device cannot detect the “white coat hypertension” syndrome [5], on top of that the traditional device to measure BP uses techniques such as oscillometric [6] and auscultation [7], which are useful for infrequent measurements but not continuous. In addition, even when the patient tried to measure BP more frequently using these techniques, the amount of discomfort and inconvenience makes it unsuccessful [8]. The proper way to get a continuous cuffless blood pressure measurement with gold-standard accuracy is catheterization which requires a medical facility with professional intervention [9,10]. So, finding a solution for the frequent measurement of blood pressure during daily life in a cuffless and noninvasive manner

remains a topic of research. The goal is to develop frequent BP measurement devices that are non-invasive, inexpensive, cuffless, wearable, and convenient.

Over the past few years pulse transit time, which is the time it takes for the pressure wave to go from one area of the body to another, has attracted a good amount of attention as a characteristic of blood flow pattern behavior in the artery to measure blood pressure [11–17]. Using photoplethysmography (PPG) and electrocardiography (ECG), the PAT can be measured. However, PEP can be influenced by different factors such as stress, age, emotion, movement, etc. In addition, it was found that PAT does not similarly correlate with DBP as with SBP [18,19]. Blood is pushed through the vessels when the heart ventricular chamber contracts maintaining expansion and contraction [20]. This regular activity impacts significantly the elastic nature of vessel walls and over time the elastic nature of vessel walls naturally degrades which in turn affects the pressure of the blood as well as velocity of pulse wave [21]. However, the accuracy is not consistent because, due to factors such as age, diet and stress, etc., the elasticity varies between individuals [22–24]. Along with PTT and PAT, the use of only a single sensor such as PPG also gained large research momentum recently. Due to being optical and very inexpensive, it became very popular among researchers to extract a different number of features from PPG waves and use formulas motivated by fluid dynamics or machine learning to come up with a model to measure blood pressure [25–30]. Volumetric change in the peripheral artery and blood pressure change is correlated [31]. This characteristic is also being used along with some other features from the PPG wave to make a model for BP measurements through different machine learning algorithms [32–36]. These techniques heavily depend on the accurate acquisition of PPG signals, on top of that there is yet to be a general formula or technique which works across different databases with equal accuracy.

Improvement in digital signal processing through the use of filters made it possible to easily process biomedical signals [37–45]. The motion artifacts, due to movement, and the heavy deviation of results, due to change in datasets, still pose a significant challenge [33] in adopting existing data-driven techniques. Evaluation with individual or another new dataset their performance fails to uphold with an acceptable accuracy range that is given by the Association for the Advancement of Medical Instrumentation (AAMI) [46]. In solving classification and pattern recognition deep learning is now a very effective and promising tool. Deep learning has the advantage of learning the features directly from the raw signals and being applied in different fields such as image recognition, speech recognition, and object identification, etc., [47–50]. A large amount of medical data is being mined for diagnostic values such as cancer detection, heart beat classification, heart disease detection brain tumor recognition, etc., [51–53]. Deep learning already started to play a vital role in research related to health states, detecting diseases, performing diagnoses, and taking preventive measures. Liang et al. proposed blood pressure measurement technique using wavelet transformation and CNN from PPG signal. The accuracy found from that experiment was 82.9% [54]. Another study adopted a model that consisted of artificial neural network in lower level for temporal feature extraction and LSTM layers in upper layer for time domain variation in ANN layer features [55]. Unfortunately, the experiment with only 39 subjects showed greater accuracy but they have not replicated it over larger datasets. Shimazaki et al. proposed a different experiment [39] using a three-layer auto-encoder to generate the features automatically and estimate BP [56]. Along with PPG, speed of PPG, derivative of PPG and health data such as age, height, weight, sex, drug presence, and pulse rate were used as inputs. The resulted deviation gold standard was 11.86 which was still quite high. Eom et al. proposed CNN-based BP estimation. Due to a very small population number (only 15), the small standard deviation is not a significant improvement [57]. Athaya and Choi proposed [58] a PPG-based deep learning technique (U net) to measure blood pressure with a significantly small standard deviation, but the population was only 100. Aguiree et al. proposed RNN encoder-decoder architecture whose performance showed large estimation errors and standard deviation of 7.32 and 15.67 for DBP and SBP [59]. All of these aforementioned experiments were done during the last three

years and considering their performance it is evident that the cases which achieved better accuracy or small standard deviation have smaller datasets or populations, and when the larger datasets or different individuals from the original dataset were used, they cannot sustain the same performance [60]. Recently, researchers used deep learning techniques to measure BP. The performance showed better accuracy in contemporary methods while using the same dataset for training and testing. Still, several shortcomings remain; first, the testing was done using the same dataset, so the higher accuracies do not carry over to new datasets or individuals. Second, in rare cases when they have chosen to test the technique with separate datasets, the size of those datasets was very small. Third, since deep learning takes care of feature extraction and selection without human intervention, it is always difficult to explain the relationship between any feature and outcome [57,61–65].

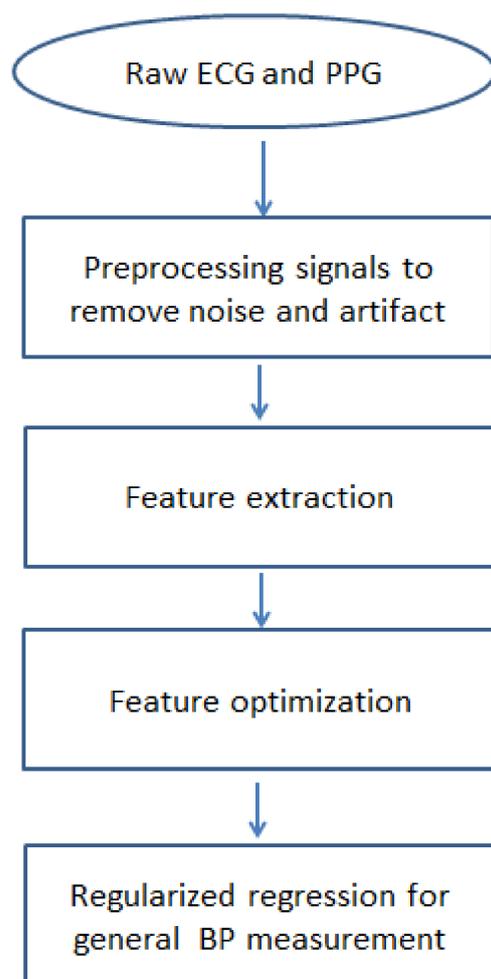
To perform feature selection and optimization, the most common and popular choice is the use of Pearson's correlation coefficient. It results in a very accurate correlation whenever the features are linear or very close to being linearly related. However, when variables are nonlinearly related, Pearson's correlation coefficient is calculated using linear approximation between such variables. Based on how nonlinear the relation is, the Pearson coefficient may or may not provide an accurate correlation value. If from beforehand it is known that the relation among some of the features is highly nonlinear, Pearson's coefficient may provide the wrong feature selection [66]. This could impact the accuracy of the BP measurements. A potential solution would be to calculate the correlation to measure independence between features irrespective of their linear or nonlinear relationship. The relationship among physiological measurements, such as the heart's electrical activity, pulse wave, and blood circulation through arteries, etc., are highly nonlinear [67–69]. For example, blood pressure and artery diameter or stiffness are related in a strong nonlinear manner and any sort of linear approximation would result in a much smaller correlation value compared to their existing nonlinear relation between feature and target variable [67]. The nonlinear dependencies decrease the chance of reducing the dataset dimensions since linear techniques or linear approximation may fail to sufficiently identify them [70]. In addition, extreme outliers or skewed distribution of data, which are common in physiological datasets, may negatively impact the performance of linear methods [70]. In contrast to the only linear model, the nonlinear models based on both linear and nonlinear dependencies show better feature selection performance with more stability and are less affected by including or excluding variables from the dataset [71,72]. For example, research was done by extending the linear correlation to an alpha-grade monomial relation who maximizes the correlation based on changing the value of alpha to which type of nonlinear relationship data exhibit [71]. Their study proved that the proposed alpha-grade correlation coefficient overcame the common drawback of statistical measure that evaluates the significance of relationship between two variables based on their linearity. These studies have provided the basis for the potential improvement of feature selection techniques for physiological data. The expectation is it will solve all aforementioned shortcomings by considering the correlation between both linearly and nonlinearly related features from vital signs, getting rid of bias towards high cardinality features and allowing total correlation in place of only pairwise.

Based on the literature survey it is evident that there is a need for a reliable and accurate cuffless blood pressure measurement system that can be used frequently by a person at his/her convenience. In this paper, a cuffless BP measurement technique that accounts for linearly and nonlinearly dependent feature with blood pressure is proposed. For the first time, a hybrid feature selection algorithm based on multivariable symmetric uncertainty (MSU) is used to optimize the number of relevant features and reduce the number of redundant and nonlinearly related to BP. The MSU-centered fast correlation and ReliefF feature selection techniques can identify dependence among features irrespective of whether they are linearly or nonlinearly related compared to the popular Pearson's coefficient which only deals with linearly dependent features. The use of the mentioned hybrid feature selection algorithm is expected to give more accurate measurements than

what is currently available in the state of the art. The rest of the paper is organized as follows: Section 2 focuses on the proposed technique for BP measurement using biomedical signal, Section 3 discusses data analysis followed by Section 4 with the result and performance of the technique using an online dataset and finally, Section 5 concludes the paper and discusses future work

## 2. Materials and Methods

The significant issues which were addressed in the technique are: first, to select the biomedical signals and features that have noticeable biological impacts on change in blood pressure; second, using symmetric uncertainty to perform a fusion of the fast correlation and Relief algorithms to optimize feature selection; third, to use regularized regression to develop a relationship between the extracted features and BP measurements. As the primary features selection is done considering their biological significance and the feature optimization algorithm is performed based on total correlation rather than just pairwise, this technique is expected to result in a smaller number of predominant features. The penalty-based regularization technique used in this study solves the problem of the unbalanced result of bias-variance found in ordinary least square regression. In Figure 1, a flowchart shows the steps of the proposed measurement technique. In subsequent subsections, these steps will be described in detail.



**Figure 1.** Diagram for model used for cuffless BP measurement.

When the raw signal is used for this kind of study the best way to proceed with that is to annotate them into acceptable and unacceptable signals. The recording of the measurement of vital signs can deviate from the ideal case due to so many reasons other

than the physical deterioration of patients. So, without annotation, it is hard to differentiate between acceptable and unacceptable signals. In this study, annotation was not performed, and we tried to find the deviated signals by using manual inspection. Ideally using an index such as signal-to-noise ratio or some other statistical parameters can be used to use only acceptable signals based on a preferred threshold. A basic classifier can be used to find out which index would perform the best. Another limitation of this kind of study is the choice of location from where the vital signs were recorded or measured. Since the dataset used in our research was from the MIMIC-II dataset, which is data acquired from hospital settings from ICU patients, the physical locations of data acquisition from patients were predefined. Study [73] shows wrist PPG is difficult to measure accurately which results in an error in peak detection; compared to that, head and finger is the much better choice. Since the distance travel by the pulse originated at the heart plays a significant part in calculating blood pressure, the best case would be to use multiple places to measure PPG signals to validate each other's data and make the measurement of pulse transit time more reliable.

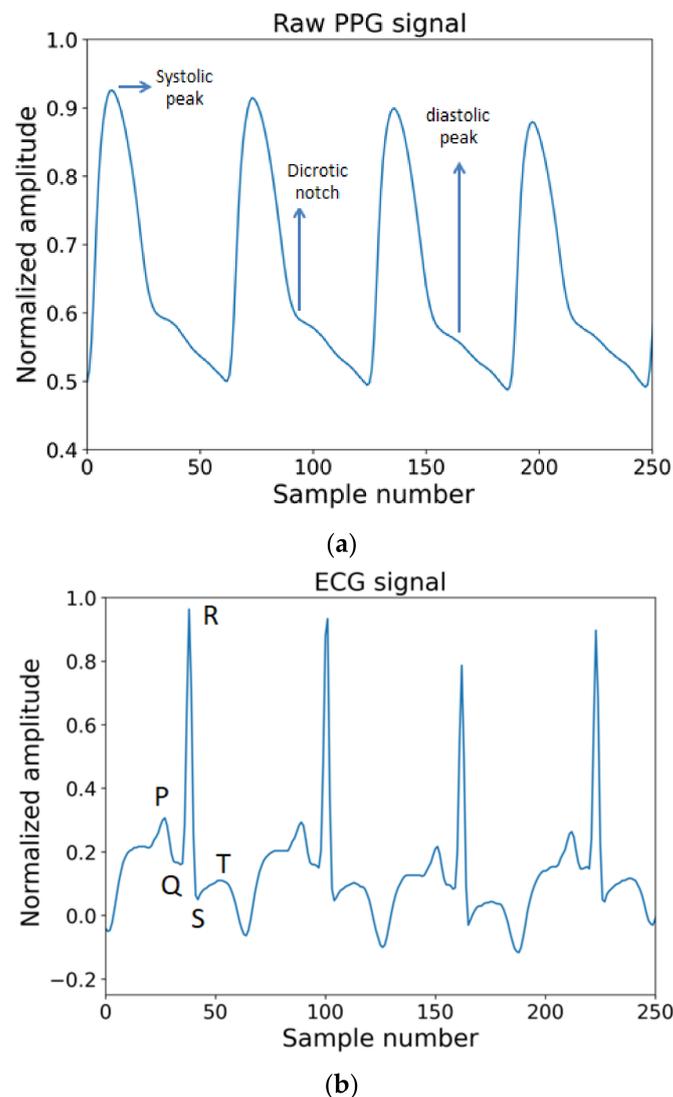
Different lead information of ECG records the electrical activity of the heart from different directions and orientations. Ideally, all 12-lead information is the best way to select any feature necessary from ECG but to make the data acquisition inexpensive and the calculation simple it is hard to accommodate more than one or a few leads at the same time. It will become cumbersome for the patient to measure the signal as well; on the other hand, not having information from all the leads deprives them of having some necessary information or the ability to cross-check any specific deviation in ECG wave. This study put focus on making the system as simple as possible while keeping accuracy and feature number optimized, so only lead information from lead II was used. The use of more lead information may help to bring forth more features that are dependent on consecutive lead information. The intrinsic stiffness of the carotid artery was only shown to increase for young patients irrespective of the change in BP but for the elderly, it is the opposite [74]. In addition, a continuous increase in BP enhances increases in vascular thickness and structural stiffness, and load on the arterial wall [75]. Damage to either small or large arteries affects the rise of central BP by increasing more wave reflections during blood flow [76]. All this information points directly to the necessity of accurately measuring the change in arterial stiffness as it changes due to the physical changes of patients, so measuring the changes in arterial stiffness using a separate sensor would be ideal.

The inclusion of statistical and spectral features along with morphological features has both positive and negative aspects for measuring blood pressure. On the positive side, more features would increase the chance of getting more relevant features in the final feature list as well as have the advantage of observing the change in features from another perspective such as statistical or frequency wise. On the negative side, the use of statistical features carries the risk of misinformation due to the improper segmentation of signals; also, lack of information about how the frequency components are directly related to the change in blood flow pose a risk of either accepting a frequency that has a minimum effect or neglecting one with higher impact. Another common limitation of this kind of study is the lack of a centrally managed dataset to test the produced model. The best way to test a trained model is to train with one standard database and test with another standard database, but unfortunately, the unavailability of biomedical data collected reliably for a longer period with high-quality raw data is still difficult to find from several sources. These drawbacks create several bottlenecks: first, people are forced to acquire data by themselves and with a limited number of data it is quite impossible to bring forth any robust model; second, even if one's model shows very high accuracy, it could fail to hold similar accuracy when another dataset is used to validate the performance. Since the datasets are differently standardized, it is very difficult to conclude whether the difference in performance came due to a difference in dataset or technique.

### 2.1. Selection of Vital Sign

The pressure on the walls due to blood flowing inside the vessels is proportional to the force that the heart uses to pump blood; this pressure is called blood pressure. The flow of blood is proportional to the pressure gradient which is the force that pushes the blood along the vessels and is inversely proportional to the resistance against the flow. From the aorta and throughout the cardiac cycle, the pressure is highest but when the large vein brings the blood back to the heart it becomes lowest and is around 0 in the vena cava area. These facts indicate that the pressure gradient is proportional to the blood pressure throughout the heart cycle [77].

The most commonly known reason behind hypertension is a build-up of plaque in arterial walls and a decrease in the elasticity as well as the radius of arteries requiring more force from the heart to pump blood and higher pressure on the vessel wall. This plaque build-up impacts patients' health in two ways. First, it progressively worsens the health condition by obstructing blood flow to different important organs of the body and second, it brings significant cardiovascular disorders. This knowledge guided us to choose appropriate vital signs which reflect changes in blood pressure. As explained in what follows, the ECG and PPG waves are the proper biomedical signals to be used to perform cuffless BP measurements. Consider Figure 2a,b, where PPG and ECG waves are shown.



**Figure 2.** (a) A single PPG signal. (b) Electrical activity pattern from ECG wave.

In the ECG wave, when the R peak in the QRS complex of the ECG wave occurs, that means the left ventricle forces the blood out to the aorta. Now when this rush of blood flow reaches the place where the PPG wave was measured, the time interval between the R peak and corresponding to any significant points in PPG wave, such as systolic peak, diastolic peak, point with the highest slope, diastolic notch, or start point of systolic phase, etc., can be measured. These time intervals have a highly correlated relation with changes in blood pressure [78]. In a study using the Windkessel model simulation, a trend line was obtained by measuring systolic rise time (SRT) and diastolic fall time (DFT) for two different compliances where compliance is the ability of the vessel to extend or be elastic enough to respond to an increase in pressure due to a larger amount of blood [78]. Blood pressure is predominantly affected by the compliance of the vessel wall. The compliance works in inverse order compared to tension or stiffness in the artery walls. The behavior from simulation also resembles the behavior of vascular change such as an increase in compliance and decrease in DFT peak [78]. The reason is when artery blood pressure increases at the same time the stiffness or tension in vessel increases which in turn decreases the compliance. With the opposite change in blood pressure, the action reverses in terms of compliance too.

## 2.2. Pre-Processing Signals

Contamination comes in the ECG wave from noise and artifacts, and the goal of pre-processing step is to reduce them as much as possible keeping the integrity of the original signal [79]. The noise and artifacts are power line interference (50–60 Hz), base-band wanderers (around 0.3 Hz), power line interference, motion artifacts, and noise from nearby electronic devices, etc. The pre-processing stage starts with a filtering block to delete the artifacts from the ECG signal [80]. An example of noisy and noise-free ECG signals is depicted in Figure 3. Typically, a band stop filter and an FIR filter of order 50 is sufficient to get rid of noise with cut-off frequencies of 0.5 Hz and 100 Hz [81]. The output of the FIR filter is being passed through a moving average filter to smooth the signal and remove the unusual momentary spike. Unlike the low pass filter, the high pass filter does not attenuate much of the signal but suffers from phase shifts [82]. The high pass filter removes the DC offset which comes from electrode/gel/body interference and is used to remove the baseband wanderer.

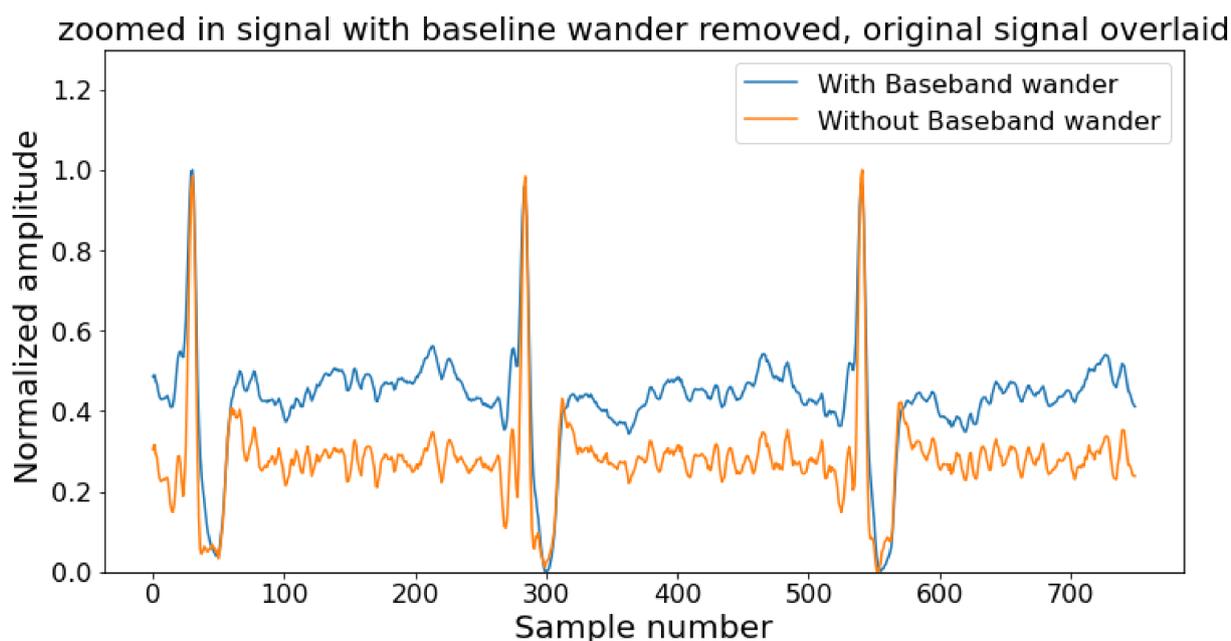
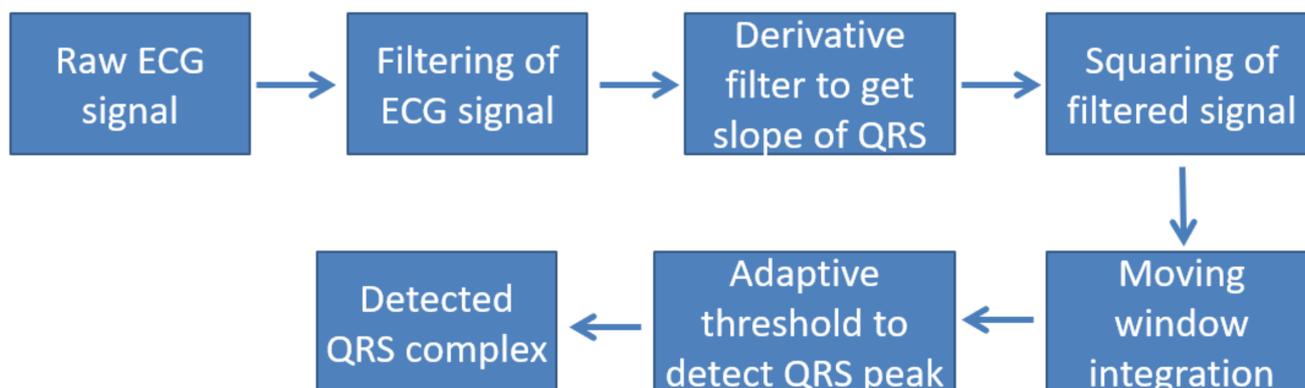


Figure 3. Example of normal noise-free and noisy signal.

Normalization is another step in the pre-processing stage which helps us to make comparisons among data from different patient at the same time. One of the most important pre-processing techniques is the use of the Pan and Tompkins algorithm [83] which includes the differentiation, squaring, and finding of the peak. In Figure 4, the flowchart for the Pan and Tompkins algorithm is depicted.



**Figure 4.** Block diagram of Pan–Tompkin’s algorithm [83].

PPG and ABP waves have similarities in terms of morphological aspects at least; FFT was used to remove high frequency noise. In lower frequency regions, a segment from 0.15 to 0.4 Hz, which is a consequence of respiratory effect, and another segment from 0.01 to 0.15 Hz vascular resistance, are divided [84]. Although in numbers there is coherence between the lower frequency region of PPG and BP, the mechanism or rationale behind the coherence is still debated among researchers [84]. Change in respiration rhythm is reflected in the change in stroke volume which in turn affects PPG wave pattern. Although researchers are not yet agreed on how much respiration affects pulse wave, it is agreed that the change is very minor when only amplitude and slopes are in question. The ideal way is to reduce or nullify the impact of respiration on PPG wave pattern by leveling the PPG wave without the envelope created by respiration.

The main peak of each PPG cardiac cycle was detected through Daubechies 3 wavelet. The DB3 is similar to the changes in PPG wave. The common peaks and intervals in PPG waves are systolic peak, diastolic peak, systolic time, diastolic time, and dicrotic notch.

The derivatives of PPG waves are very useful for interpreting original PPG signals. In Figure 5, the original and the first derivative and Figure 6 shows a PPG signal and its second derivative. In Figure 5, the significant points in the first derivative were shown as  $\Delta T$  and crest time (CT). The  $\Delta T$  is the time interval between the first derivatives of PPG where the signal is going to zero from a positive direction to negative directions. The CT is defined as the time interval from the minimum of the PPG wave to its maximum peak.  $\Delta T$  and CT are closely related to any abnormality related to a heart condition [85]. In Figure 6, the significant points are shown as points a, b, c, d, e. According to a previous study, the ratios, such as  $b/a$ ,  $c/a$ ,  $e/a$ , and  $(b-e)/a$ , etc., strongly related to the change in arterial stiffness [86]. The points a~e and the time interval between different points of ECG and PPG are used as features to measure blood pressure.

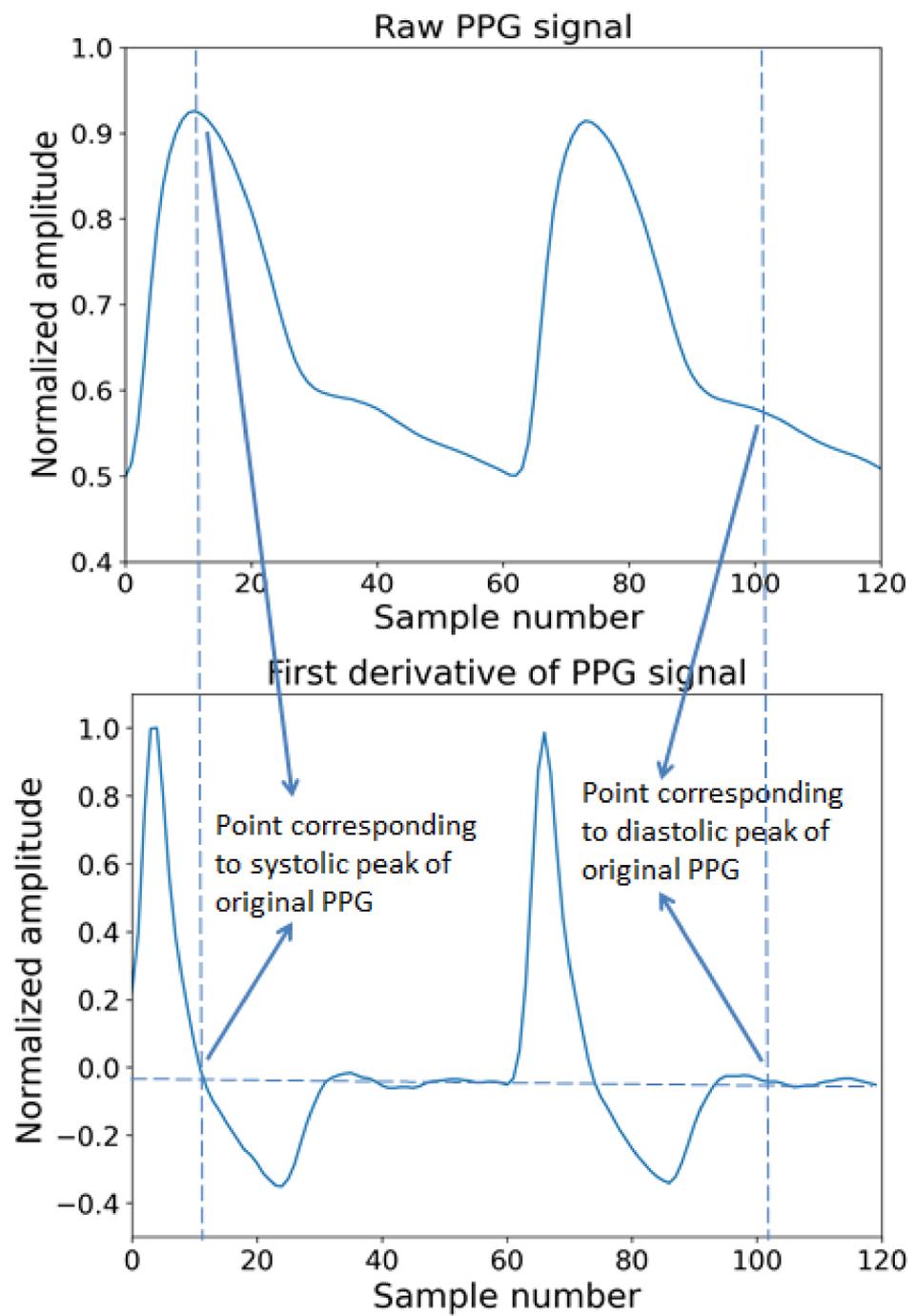
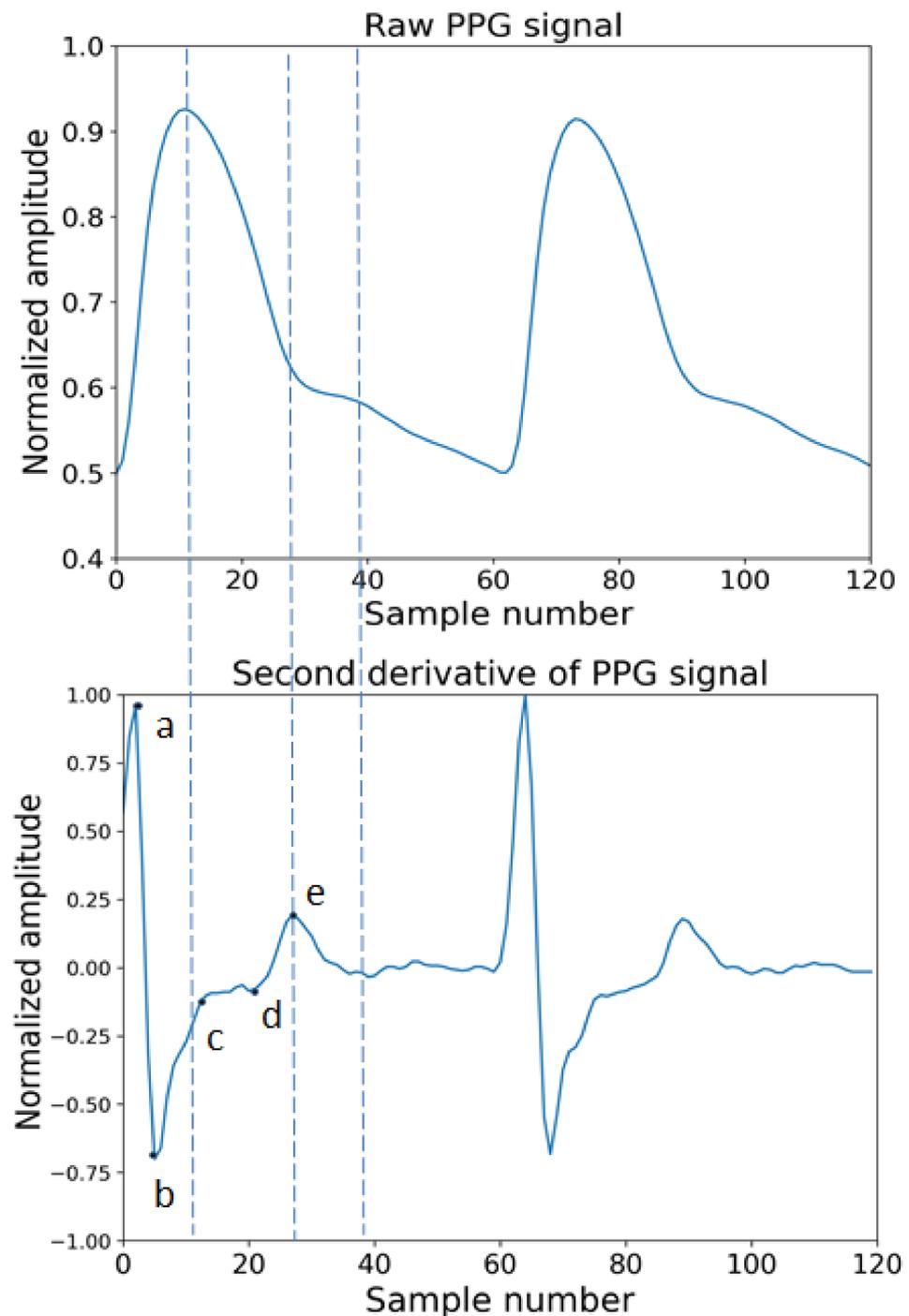


Figure 5. Original raw PPG wave and first derivative of PPG wave.

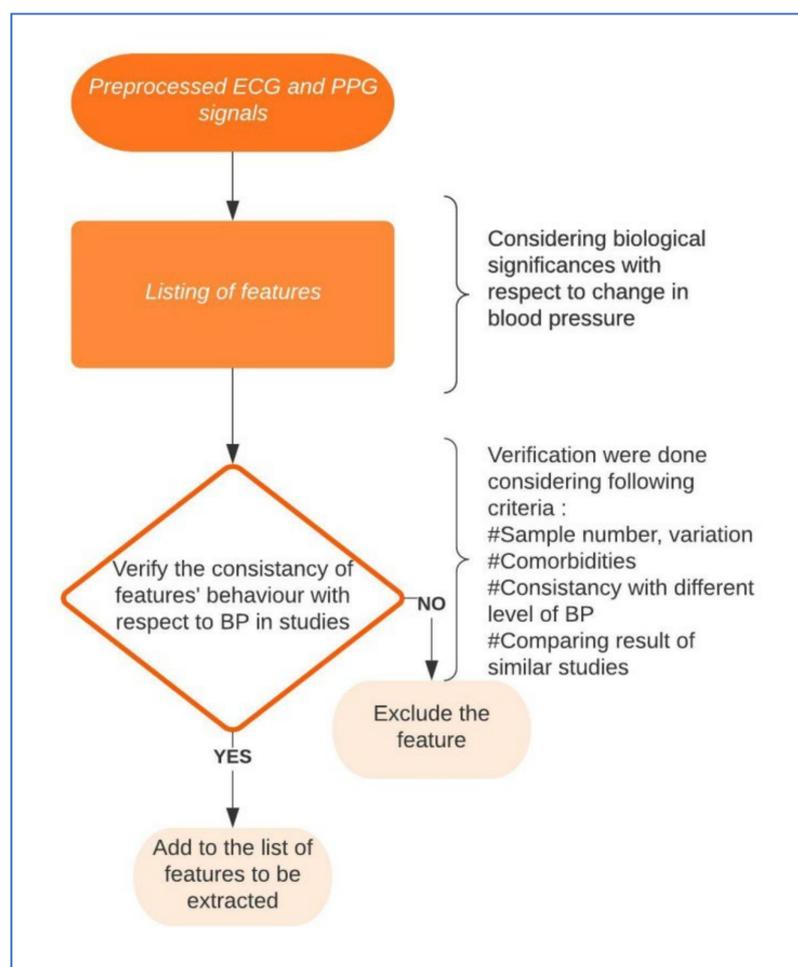


**Figure 6.** Original raw PPG wave and second derivative of PPG wave.

### 2.3. Feature Extraction

The whole process of finalizing the features to be extracted is shown in Figure 7. First, the morphological features from ECG will be discussed which are relevant to blood pressure measurement. The main challenge with ECG features is that the morphological changes in ECG for a hypertensive patient can either change for a short time period at the onset of high pressure or the change can sustain for a longer period [87]. The characteristics of the P wave from an ECG envelope, such as P wave area, amplitude, minimum duration, amount of dispersion, etc., have shown some relation with a change in blood pressure [88–90]. The length of PR interval increases with the increase of blood pressure [89–92]. The QT interval is the duration between Q waves until the end of the T wave, the characteristics

which were found related to BP changes were: QT duration maximum, minimum, and QT dispersion [91]. The fragmented QRS complex is found more often with high blood pressure patients, also with a higher amplitude of the R and S wave. In the case of the T wave, the features are the same but the P wave was found to be related in some way to a change in blood pressure [90,93–95]. In previous studies, along with other discoveries, there were some shortcomings. For example, in some studies [91,92,94], the BMIs were not well spread out, some have patients mostly of a young age, some have limited patients with specific prior diagnosis, and several studies did not report any comorbidities or confounders, etc., [90,95]. These limitations require further study with these features before finalizing a specific list. Till today, there is no clear list of features from the ECG wave that are optimized in identifying hypertension reliably [96].



**Figure 7.** Diagram with the steps to finalize the features extraction process.

In the PPG wave, the first phase is related to the systolic period and the second phase is related to the diastolic period. Systolic amplitude is found to be suitable for measuring pulse arrival time [97]. The pulse width is the time duration for the whole pulse.

Researchers use different parts of pulse width for their experiment, such as using a half or quarter of the amplitude of systolic peak for their experiments [98]. Since the pulse width varies with the vascular resistance using the dicrotic notch, the area under the wave can be divided into two parts. The ratio of those parts indicates a strong relationship with total peripheral resistance, which is the force necessary to keep the blood flow from the aorta to the venous exit into the auricles [99]. An increase in total peripheral resistance will increase the blood pressure immediately. A recent study concluded that the correlation

between PPG waveform and ABP waveform is very low. Therefore, the ABP cannot be replaced by the PPG wave for BP measurement.

They used synchronized signals, and a strong correlation indicated that their morphological changes are related and depend on each other; at the same time as the ABP increases, some different phases between the signals became visible. These changes have not been studied extensively yet so, although the morphological relationship between the signals shows a strong correlation with the increase of ABP, the change in correlation needs to be studied [100].

#### 2.4. Feature Optimization and Selection

To identify the significant features from the available feature list, the feature selection and optimization algorithm need to be implemented. With the use of only relevant features the classification or data mining model ends up providing results with better accuracy and reducing computational complexity. A fundamental problem with any data is to approximate a relationship between input  $X = (x_1, x_2, x_3 \dots x_N)$  and output  $Y = (y_1, y_2, y_3, \dots y_N)$ . Sometimes, output  $Y$  cannot be determined by using all features but only a subset of features. If all the features including the irrelevant features were used there are a few problems that might occur. First, the irrelevant features will require more computational cost (sometimes polynomial greater) while doing any prediction. Secondly, the irrelevant features may produce overfitting which will fail the model when it is tested with any independent data from which it was trained on. In addition, when the mode is too simple and includes irrelevant features as well, under fitting occurs. That means it tends to have less variance in the prediction but more bias towards wrong outcomes [101]. The major advantages of optimal feature selection are: First, simpler models are simple to interpret. As the number of variables increases, the model gets more and more difficult to interpret. Second, the small number of features results in quicker training and prediction time. Third, redundant information from different features creates a false interpretation of higher accuracy but fails in the case of testing with data from a different dataset.

The filter-based method relies on feature characteristics and is favorable for the model which requires quicker selection. The wrapper method uses a machine-learning algorithm to select the best features in the subset and it is more computationally extensive than the filter-based method [102]. The embedded method is an iterative approach that extracts features that contribute most to the training for a particular iteration. In our study, the filter-based approach was primarily selected based on the following reasons. Unlike the wrapper-based method, the filter method does not incorporate any machine learning model to determine the relevancy of any feature. The filter method is much faster compared to the wrapper method because they do not require any training phase, so with a larger dataset, the difference in computation becomes quite large. Unlike the wrapper method, the filter-based method does not depend on a heuristic search algorithm. In the hybrid approach, the combination of filter and wrapper does not integrate well which may result in lower classification accuracy [103]. Popular filter-based feature selection methods use different techniques to differentiate between relevant and irrelevant features. For example: Euclidean/Manhattan distances between feature variables project dependence or correlation within features as well as between feature and target variable. An example of using Euclidean distance between features is the relief algorithm [104]. To better understand the aforementioned “distance” feature, Figure 8 is provided to introduce what is a nearest ‘hit’ or ‘miss’.

For a given target instance, which is shown in Figure 8, when it is compared with another instance with same class variable it is called a ‘hit’, and the other way around is called a ‘miss’. For a nearest hit and sample feature, such as A, B, C in Figure 8, when the values of the feature are different in those cases, the corresponding feature weight is decreased by  $1/(\text{total feature number})$ . Similarly, in the case of a nearest miss for same scenario, weight for the corresponding feature is increased by same margin. An improved version of this technique involves calculating the distance function [105] of a specific feature

(for example B) using the ratio of difference of value of that feature between target instance (TI) and the nearest hit/miss and the difference between the maximum of B and minimum of B.

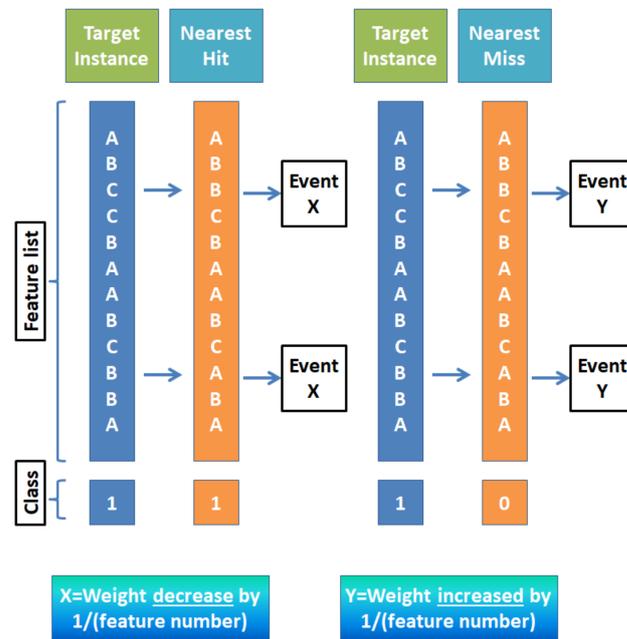


Figure 8. An example of feature weight update procedure based on distance target instance.

The distance function is also used to find the nearest neighbors by calculating distance between different instances. The total distance is basically the summation of all the distance gathered from all features. An example of using dependence is Pearson or Chi-squared, where dependence or independence between two variables is measured [106]. Mutual information or information gain where information is shared between features or contributions of information of a feature are measured to differentiate between features [107].

In this paper, a combination of two feature selection methods is used. The first method is the fast correlation (FC) where it checks the correlation of each feature and relevance with the target variable. In 2021, it was discovered that FC has a problem of sometimes considering relevant features as redundant and ignoring them [108]. So, solely depending on FC cannot be the optimum choice. That is why the second method which is a ReliefF-based algorithm has been added in this work. It uses Manhattan distance instead of Euclidean distance in the Relief algorithm to differentiate between the same or different types of feature classes. ReliefF considers conditional dependencies or interactions between features which usually get ignored by the FC algorithm [109]. In both, the method of symmetric uncertainty was used as a measure of goodness for the features which is entropy-based correlation while correcting the bias of information gain. The reason behind choosing this over Pearson’s coefficient is the option to encompass the non-linear relationships between features since most physiological measurements are nonlinear [110]. In addition, the Pearson algorithm has very little robustness against outliers [111].

#### 2.4.1. Symmetric Uncertainty

To measure the feature effectiveness for classification, the feature has to be relevant to the class and at the same time not be redundant to any other relevant features. In this work, correlation is used as the criteria to select features. When the correlation between two features is small, but correlation with the target variable is higher than a certain threshold

compared to another relevant feature, such a feature is selected as effective for that specific classification task. The linear correlation coefficient is as follows:

$$r = \frac{\sum_i (x_i - avg(x_i))(y_i - avg(y_i))}{\sqrt{\sum_i (x_i - avg(x_i))^2} \sqrt{\sum_i (y_i - avg(y_i))^2}} \tag{1}$$

The mean  $r$  lies between  $-1$  and  $1$  and represents the full correlation range between  $X$  and  $Y$ . When they are fully independent then the value of  $r$  will be zero. Correlation analysis removes the features which have a near-zero linear correlation with the class at the same time it reduces the redundancy between features considered linearly correlated. The limitation of linear correlation analysis is the assumption of linearity between features, which is not true for all features. In addition, all the features must only be in the numeric form to be considered for correlation analysis. The linear correlation coefficient assumes a linear relationship between variables even in cases where the relationship is quite non-linear, so just by getting a low or near-zero Pearson’s correlation coefficient it is difficult to correctly conclude that they are independent of each other.

To solve these shortcomings, a potential solution is to choose a correlation measure based on entropy and use mutual information (MI), which is the reduction in uncertainty of one variable while observing the other variable. If we compare how the joint distribution of the variables differs from the multiplication of their marginal distribution, it can be concluded that they have dependence whether they are linear or nonlinear [112]. A discrete random variable  $X$  with values  $\{x_1, \dots, x_k\}$ . The entropy for the variable  $X$  is  $H$  which measures the uncertainty of the prediction of  $X$  [113]. So, the entropy is defined as:

$$H(X) = - \sum_i P(x_i) \log_2(P(x_i)) \tag{2}$$

For  $x_i$ , the prior probability is  $p(x_i)$ , and the entropy  $H(X)$ .  $H(X|Y)$  can be expressed as follows [114]:

$$H(X|Y) = - \sum_j [ P(x_i) ] \sum_i P(x_i|y_j) \log_2(P(x_i|y_j)) \tag{3}$$

Now the mutual information can be introduced which measures the reduction in the uncertainty of  $X$ .

$$MI(X|Y) = H(X) - H \tag{4}$$

In case of  $X$  and  $Y$ , they are independent of each other,  $H(X,Y) = H(X) + H(Y)$ ,  $M(X|Y) = 0$  and  $H(Y|X) = H(Y)$ . In contrast, when  $X$  and  $Y$  are fully correlated, the joint entropy  $H(X|Y) = 0$  and mutual information  $M(X|Y) = H(X)$ . Although symmetry is a desirable property when we want to consider it to be an effective measure for features, MI or information gain (IG) are biased towards high cardinality features. So, MI or IG needs to be normalized using the entropies of the features  $H(X)$  and  $H(Y)$  to compensate for such bias. The resultant correlation measure is called symmetrical uncertainty (SU) [115], and it is given by

$$SU(X,Y) = 2 * \frac{Mutual\ information}{H(X) + H(Y)} \tag{5}$$

The range of  $SU$  is between  $0$  and  $1$ , where the value  $0$  indicates that  $X$  and  $Y$  are independent. Even with  $SU$ , there exists a limitation of computing more than two variables. So, an extension is necessary for the multivariable cases to detect interactions among different sets of variables.

Before moving towards multivariable symmetric uncertainty (MSU), let us first discuss the concept of total correlation [116]. For a given vector of variables  $X$ , the joint entropy for  $n$  number of variables is  $H(X_{1:n})$ , given by [114]

$$H(X_{1:n}) = - \sum_{x_1} \dots \sum_{x_n} P(x_1, \dots, x_n) \log_2[P(x_1 \dots, x_n)] \tag{6}$$

For  $n$  number of variables, the mutual information can be expressed as a reduction of the uncertainty among one or more variables by knowing another. It is a comprehensive measure of independence, which means when the variables are independent then joint entropy equals the sum of their marginal entropies. When it is not equal then some sort of dependency is present. That is why it is better to consider the linear and nonlinear dependencies compared to linear correlation [117]. For  $n$  number of variables, the total correlation is shown in Equation (7). Here, total correlation does not indicate only linear correlation but linear and nonlinear dependencies, thus giving total mutual information:

$$C(X_{1:n}) = \sum_{i=1}^n H(X_i) - H(X_{i:n}) \tag{7}$$

If we use  $n = 2$ , then the total correlation will turn into information gain or mutual information between the two variables.

$$\begin{aligned} C(X_{1:2}) &= C(X_1, X_2) \\ &= \sum_{i=1}^2 H(X_i) - H(X_{1:2}) = H(X_1) + H(X_2) - H(X_1, X_2) \tag{8} \\ &= H(X_2) - H(X_2|X_1) = H(X_1) - H(X_1|X_2) \end{aligned}$$

$$\begin{aligned} &= H(X_1, X_2) - H(X_2|X_1) - H(X_1|X_2) \\ &= -\sum_{X_1, X_2} P(X_1, X_2) \log(X_1, X_2) + \sum_{X_1, X_2} P(X_1, X_2) \log \frac{P(X_1, X_2)}{P(X_1)} + \sum_{X_1, X_2} P(X_1, X_2) \log \frac{P(X_1, X_2)}{P(X_2)} \tag{9} \\ &= \sum_{X_1, X_2} P(X_1, X_2) \log \frac{P(X_1, X_2)}{P(X_1)P(X_2)} = I(X_1, X_2) \end{aligned}$$

Zero or close to zero value of  $C(X)$  means that all the variables are independent of each other. To transfer the concept of total correlations into multivariable symmetric uncertainty (MSU) keep the total correlation or total mutual information  $C$  between 0 and 1 and recognize that a higher value of  $MSU$  should indicate a higher correlation among variables. Equation (4) is the mutual information between two variables and for several  $n$  variables according to Equation (7), the total mutual information can be expressed as:  $\sum_{i=1}^n H(X_i) - H(X_{i:n})$ . By definition, the symmetric uncertainty is the normalized mutual information, so the denominator of Equation (5) becomes the summation of entropy of all variables which is equal to:  $\sum_{i=1}^n H(X_i)$ . The symmetric uncertainty of two variables from Equation (5) becomes the following equation as the general formula for  $MSU$  where value 2 in the case of two variables is expressed as a normalization multiplier in the case of  $MSU$ :

$$\begin{aligned} MSU(X_{1:n}) &= \text{normalization multiplier}(m) * \frac{\sum_{i=1}^n H(X_i) - H(X_{i:n})}{\sum_{i=1}^n H(X_i)} \tag{10} \\ &= m * [1 - \frac{H(X_{i:n})}{\sum_{i=1}^n H(X_i)}] \end{aligned}$$

To prove Equation (10), using the chain rule [118],

$$\begin{aligned} H(X_{i:n}) &= \sum_{i=1}^n H(X_i|X_1, \dots, X_{i-1}) = \sum_{i=1}^1 H(X_i|X_1, \dots, X_{i-1}) + \sum_{i=2}^n H(X_i|X_1, \dots, X_{i-1}) \\ &= H(X_1) + \sum_{i=2}^n H(X_i|X_1, \dots, X_{i-1}) \end{aligned}$$

This proves that  $H(X_{i:n}) \geq H(X_1)$  or in general  $H(X_{i:n}) \geq H(X_i)$ .

So, (a)  $n * H(X_{i:n}) \geq \sum_{i=1}^n H(X_i)$ .

In addition, similarly,  $H(X_i|X_1, \dots, X_{i-1}) \leq H(X_i)$ .

This result:  $H(X_1) + H(X_2) + \dots + H(X_n) \geq H(X_{i:n})$ .

Or (b)  $\sum_{i=1}^n H(X_i) \geq H(X_{i:n})$ .

Using the inequality of (a) and (b),

$$\frac{1}{n} \leq \frac{H(X_{i:n})}{\sum_{i=1}^n H(X_i)} \leq 1$$

Since the MSU in Equation (10) goes from 0 to 1, using the above right-side limit of  $\frac{H(X_{i:n})}{\sum_{i=1}^n H(X_i)}$  and definition of MSU,

$$MSU(X_{1:n}) = m * \left[ 1 - \frac{H(X_{i:n})}{\sum_{i=1}^n H(X_i)} \right] = m * \frac{n-1}{n} = 1,$$

So,  $m = \frac{n-1}{n}$  MSU can be reduced to:

$$MSU(X_{1:n}) = \left( \frac{n}{n-1} \right) * \left[ 1 - \frac{H(X_{i:n})}{\sum_{i=1}^n H(X_i)} \right] \tag{11}$$

Here the nm represents the normalization multiplier; from [119] the nm was defined as  $n/(n - 1)$ . In the case of two variables, the value of nm was 2. To prove that SU and MSU are the same measures, it can be derived that when  $n = 2$  the MSU will become SU.

$$\begin{aligned} MSU(X_{1:2}) &= \left( \frac{2}{2-1} \right) * \left[ 1 - \frac{H(X_{i:2})}{\sum_{i=1}^2 H(X_i)} \right] \\ &= 2 \frac{H(X_1)+H(X_2)-[H(X_1)+H(X_2|X_1)]}{H(X_1)+H(X_2)} \\ &= 2 \frac{H(X_2)-H(X_2|X_1)}{H(X_1)+H(X_2)} = SU(X_2, X_1) = SU(X_1, X_2) \end{aligned} \tag{12}$$

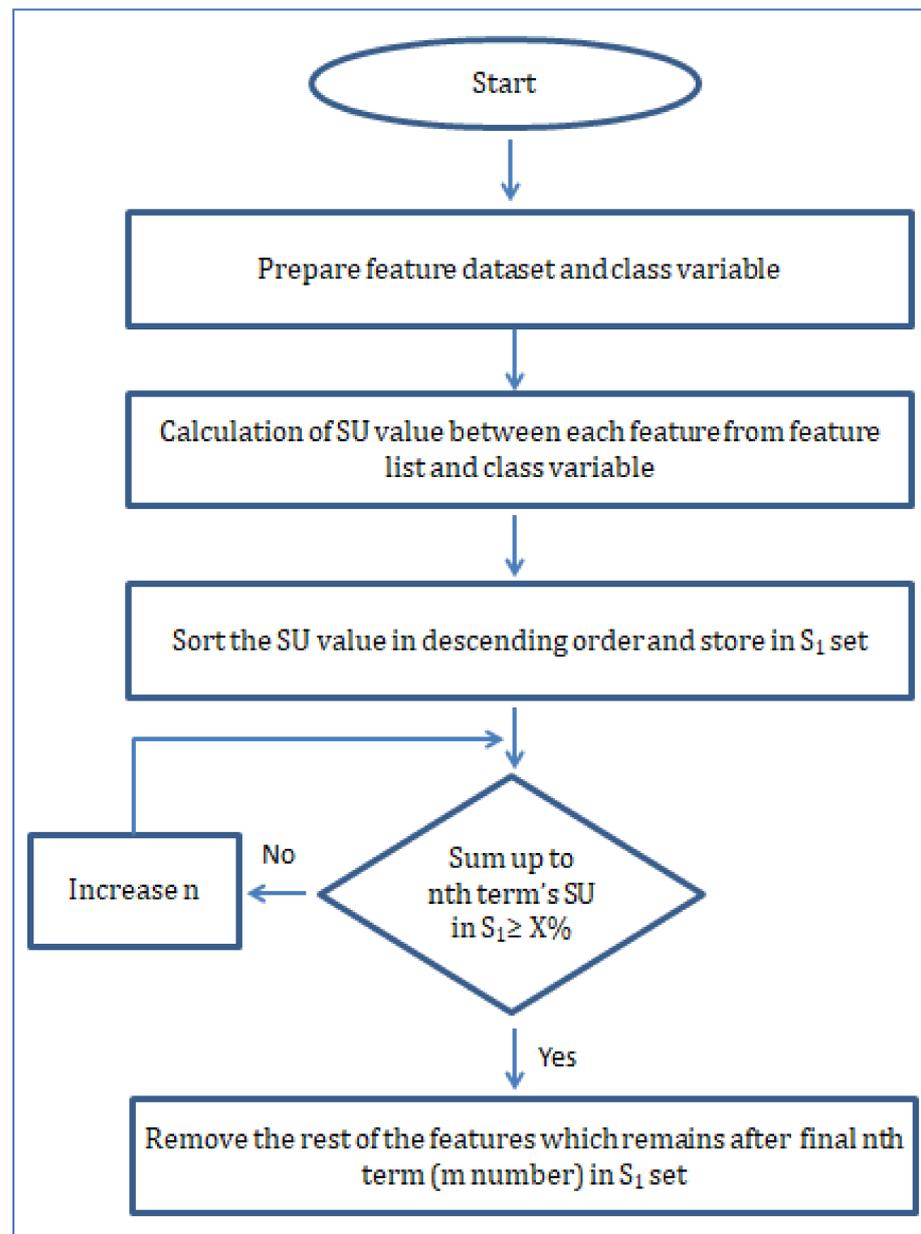
#### 2.4.2. Fast Correlation Algorithm Using MSU

In this research, the fast correlation method which was introduced in 2004 is combined with MSU to perform the calculation of feature selection. The algorithm works in two steps. In the first step, the features are sorted using MSU, and a threshold setting is used to remove less correlated features. In the second step, the comparisons among the features were done to remove the redundant features using MSU. So, this algorithm uses the accelerated pace advantage of filter and achieves higher efficiency of calculation at the same time.

Here the two steps of the algorithm are shown in two separate flowcharts. The first flowchart in Figure 9 shows how the relevance of features is measured, and in the second flowchart in Figure 10, the process of removing redundant features is shown.

Here in Figure 9, the diagram of step 1 of the algorithm has been depicted. The majority of step 1 involves calculating SU values between features and class variables to make a ranking list. Based on the predefined high threshold, the selected features are put into a set called  $S_1$  as a collection of strong relevant features with class variables. Although the  $S_1$  set includes features that show a strong relation with the class variable at the same time, it does not confirm whether there are any redundant features present or not in set  $S_1$ . For that purpose, step 2 of the algorithm is depicted in Figure 10. The main objective of step 2 is to remove redundant features from  $S_1$  and only keep the predominant features. When a feature that is ranked high in set  $S_1$  has a higher SU with the next feature in set  $S_1$  than the SU between the next feature and class variable, we can safely confirm the next feature as redundant. Using this principle, all the features in  $S_1$  are used in the same calculation pattern and marked as either redundant or not. While making the comparison using SU, a constant  $p$  is included to have a final list with a varied number of features.

Based on the MSU of the feature list in  $S'$  after each round, the value of  $p$  will be reassigned (decreased by 0.1  $p$  until  $p = 1$ ) to make the feature number in  $S'$  smaller. Finally, after achieving a specific near-zero value of MSU, this points to an  $S'$  which consists of only predominant features.



**Figure 9.** Step 1: Flowchart of fast correlation using SU value to find the relevant feature set.

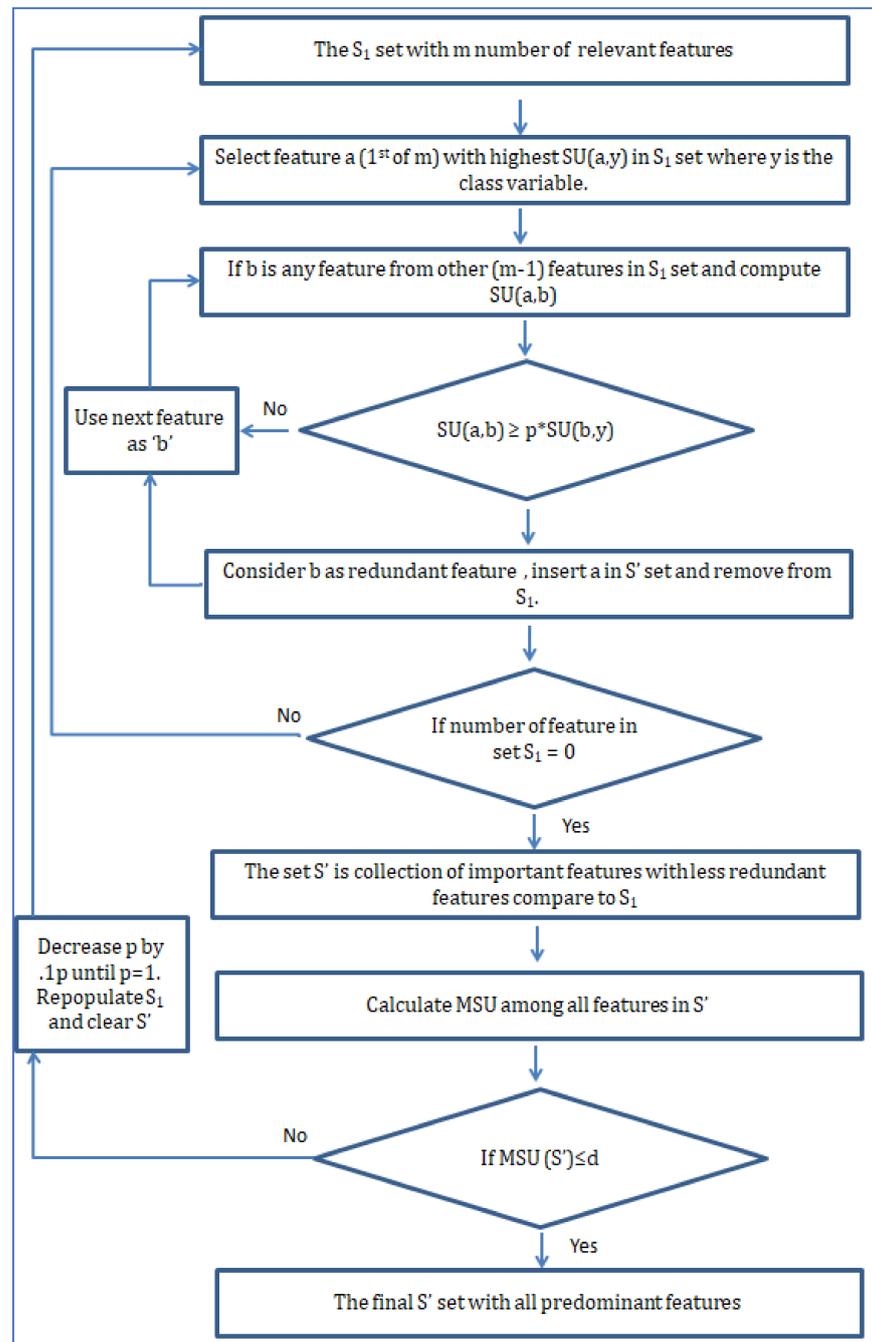


Figure 10. Step 2: Flowchart of fast correlation using SU and MSU to remove redundant features.

### 2.4.3. Relief Algorithm

Most of the feature selection algorithms assume the conditional independence of features (dependent on target variable but not on each other), these algorithms are not appropriate for solving a problem where there are many feature interactions. Relief algorithms do not assume so the quality of attributes can still be extracted with high interaction among them. The main idea is how their contribution and separation patterns behave when they are close to each other in different instances [105]. Depending on whether the random instance  $X_i$  and  $H$  result in different values for the  $F$  feature then  $F$  identifies them as belonging to the same class, but this scenario is not desirable, so the weight of  $F$  is reduced. On the other hand,  $X_i$  and  $M$  result in different values for the  $F$  feature, then  $F$  identifies them as the same class, but this scenario is desirable, so the weight of  $F$  is incremented. Let

us assume the whole instance list is  $Y_1, Y_2, Y_3, \dots, Y_a$  which are used to describe a set  $F$  of attributes  $F = 1, 2, 3, \dots, n$ , where  $n$  is the number of features. So, the function can be expressed as below for nominal features [120]:

$$diff(F, Y_1, Y_2) = \begin{cases} 0, & \text{when } value(F, Y_1) = (F, Y_1) \\ 1, & \text{otherwise} \end{cases} \tag{13}$$

For numeric features the difference function becomes as below [120]:

$$diff(F, Y_1, Y_2) = \frac{|value(F, Y_1) - value(F, Y_2)|}{Max(F) - min(F)} \tag{14}$$

Regarding limitations, research in a Relief-based algorithm has been limited to two-way interactions only [120]. It has been proved in several types of research that empirically the core performance of the Relief algorithm deteriorates as the number of irrelevant features increases, so before using the feature through a Relief-based algorithm it is beneficial to pass them through another process to decrease the number of irrelevant ones [121]. In addition, the Relief algorithm cannot work with multiclass endpoints. So, there must be an improved algorithm that can provide the advantage of Relief and solve the limitations mentioned before.

#### 2.4.4. ReliefF with MSU

ReliefF is the best-known Relief-based algorithm [104] which is the fixed algorithm variation (A to F). There are four key aspects to the improvement in the ReliefF algorithm. First, ReliefF uses the user-defined  $K$  number of neighbors and updates in the Relief algorithm [122]. Second, in ReliefD (which is also incorporated in ReliefF) the best solution to encounter missing data was proposed by equating diff function to probability that for a given feature two instances do not show the same value [123]. Third, ReliefF proposes an approach to handling multi-class target points where ReliefF calculates the prior probability from class while finding  $k$  nearest misses. Fourth, the reliability of weight estimation becomes credible as  $m$  becomes close to  $n$  [105].

The most important step in the ReliefF algorithm is after calculating the diff function, the estimation of weight for each feature, which is [124]

$$W[F]W[F] = \frac{\sum_{j=1}^k diff(F, Y_i, Y_j)}{n.k} + \sum_{C \neq class(Y_i)} \frac{p(C)}{1 - p(class(Y_i))} * \frac{\sum_{j=1}^k diff(F, Y_i, Y_j(C))}{n.k} \tag{15}$$

$Y_i$  is a randomly picked sample at time  $t_i$ ; the nearest number of neighbors of  $Y_i$  is  $k$ . Although the Relief algorithm has advantages such as being non-myopic, its ability to estimate the quality of features requires less time than the exhaustive search approach. Its main limitation is the inability to consider feature interactions for more than two features [125].

The ReliefF algorithm proposes a correlation between the calculation pattern and each fault category. The objective is to find  $K$  nearest neighbor samples from the sample set. The other way around is called near hits. In this paper, the MSU is used to create a ranking of features based on symmetric uncertainty and use the ranking for the next steps. This will reduce the computational complexity at the feature weight measurement step compared to the use of every feature in that step.

#### 2.5. Regularized Regression with a Penalty Factor

In this section, the necessity of regularization in regression technique will be discussed along with the lasso, ridge, and combined effect of those two regression techniques. Since the result of the techniques discussed in the previous subsection was to optimize the number of features, this subsection will discuss the regularized regression technique which will be implemented using those features to get a cuffless measurement method of blood pressure. The basic linear regression model predicts  $n$  observations of target variable

Y using the linear combination of m predictor variable X. If the error term is normally distributed with variation  $\sigma^2$ :

$$Y = X\beta + \epsilon \tag{16}$$

In the ordinary least square (OLS) approach, the  $\beta$  is estimated in such a way that the sum of squares of residuals is as small as possible. So, the objective is the minimizing loss function:

$$\text{minimize } \{SSE = \sum_{i=1}^n [y_i - y_i(\text{mean})]^2 = \|y - X\beta_{est}\|^2\} \tag{17}$$

In statistics, two major components to measure the characteristics of estimators or predictors are bias and variance. The bias is the difference between the estimated target variable and true output variable and bias mainly measures the accuracy of the prediction. On the other hand, variance calculates the variation or spread or uncertainty of the predictions and the unknown error variance which can be expressed as below [126]:

$$\text{error variance} = \frac{(y - X\beta_{est})(y - X\beta)}{n - m} \tag{18}$$

The expectation of OLS is low bias and variance since large values indicate poor performance of the OLS approach. The total error from the model can be categorized into three segments as follow [126]:

$$\text{Model error} = \text{error from bias} + \text{error from variance} + \text{unexplained error}(\sigma^2) \tag{19}$$

Based on whether the value for bias or variance is high or low there can be four different situations. The high value for variance and bias provides the worst prediction and, on the other hand, low value of variance and bias provides the best performance. The OLS generally provides a solution where it is unbiased, but the variance is high, so for this specific situation three problems need to be solved. They are multicollinearity, interpretability, and insufficient solution. Finally, the third issue is the balance between variance and bias (interpretability). If the model complexity and error terms are plotted, as the model complexity (number of features) increases three things happen: first, error due to bias decreases sharply; second, error due to variance increases. The total error starts high with a low number of features; it comes down to the lowest value for a certain number of features and then grows large with more features in the model. So, keeping increasing the feature number is not the solution, but the objective should be to find the optimized situation with the right balance between bias and variance. A solution for all these constraints is adding a regularization term or penalty term with OLS which will lower the variance at the cost of some bias to find the optimum solution.

When there are many predicting variables with small effects from each of them which are correlated among each other, ridge-type regression can prevent the linear regression model from showing poor performance (high variance) [127]. In ridge regression, not only does the OLS loss function get minimized, but also the size of parameter estimates gets penalized too as shown in Equation (20) [127].

$$\text{minimize } \{SSE + \text{lambda} \sum_{j=1}^p \beta_j^2\} \tag{20}$$

The lambda ( $\lambda$ ) is the regularization penalty. The lasso regression technique is used for large datasets. The change from ridge regression to lasso is the penalty term as follows:

$$\text{minimize } \{SSE + \text{lambda} \sum_{j=1}^p |\beta_j|\} \tag{21}$$

The major shortcoming of the lasso technique is oracle property and instability with high-dimension data [128]. Although both techniques use correlated predictors, they solve the multicollinearity differently. In ridge regression, the coefficients of correlated predictors

are reduced but remain like each other, and in lasso regression, the correlated predictors end up with a large value of coefficient while the rest of the predictors are zeroed. Since it is not possible to determine true optimized parameter values, there has to be a way to use the penalty term in a way to optimize the coefficient values.

To take advantage of both the ridge and lasso methods, a hybrid approach can be to use a combination of them as follows:

$$\text{minimize } \{SSE + \text{lambda}_2 * \alpha * \sum_{j=1}^p |\beta_j| + \text{lambda}_1 * \frac{1-\alpha}{2} * \sum_{j=1}^p \beta_j^2\} \quad (22)$$

Here, the value for  $\alpha$  makes the choice between ridge and lasso technique, where  $\alpha = 0$  means ridge regression and  $\alpha = 1$  means lasso regression. There is an extra quadratic term that results in a convex loss function. The inclusion of a little bias while reducing the variation is the appropriate solution. In lasso, it is a possibility that some of the features will be pushed towards zero and even totally zero, so ridge regression is more effective since it does not drastically reduce the feature number but rather takes a systemic approach of reducing feature weight. The following Figure shows the change in number and value of coefficients lambda and alpha in different situations for a sample scenario.

When alpha is equal to one, the elastic net becomes lasso regression, and the number of coefficients becomes smaller while the value of lambda gets larger. On the other hand, when the value of alpha is equal to zero, the elastic net becomes ridge regression, and the number of coefficients remains the same for different values of lambda indicating that ridge does not reduce the number of a predictor. So, when the value of alpha is between zero and one the pattern in which the value of coefficient and number of coefficient changes differs from ridge and lasso. To get the optimum number of features the implementation of cross-validation needs to be used. From the mean square error versus lambda plot, the number of the coefficient which will keep the MSE smallest for a certain value of lambda can be achieved.

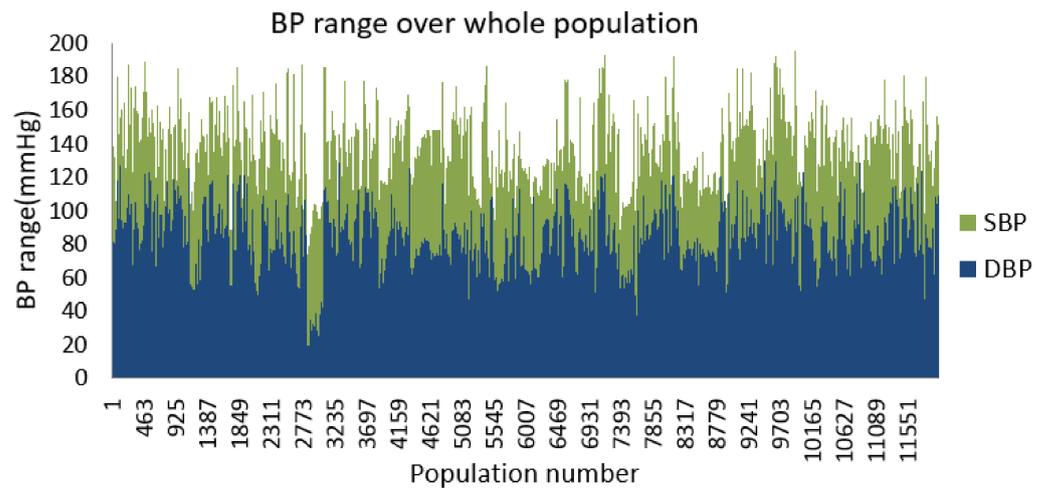
### 3. Data Analysis

In this section, the dataset which has been used in this paper is introduced and later the data analysis process will be described.

#### 3.1. Dataset

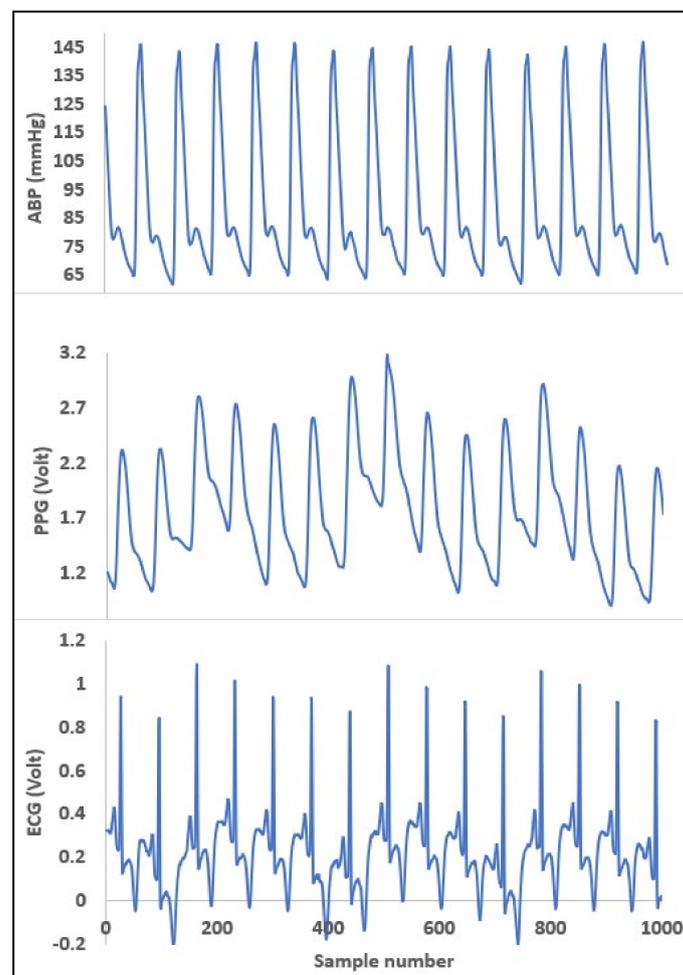
A dataset called “MIMIC II” used in this research is available on the physionet website [129]. From intensive care patients at the hospital, several vital signs were measured synchronously for a specific time frame, the vital signs which were recorded are ECG, arterial blood pressure, and fingertip pulse photoplethysmography. Between 2001 and 2008, a total of 26,870 patients were admitted in ICU. The dataset can be divided into two sections, the first one is clinical data and the second one is the physiological waveform of the measured vital signs. The waveform was collected as a raw dump file where there is some separate file for each patient and at the same time some files contain more than one patient due to the machine not being reset between patients. Those cases were separated by manually checking for the exceptional long intervals between readings. Although the initial recording of ECG was done in 12-bit precision and a higher sampling rate, later for the ease of calculation it was reduced to 7–10 bits. Later “peak picking” techniques were used to scale down the sampling rate to 125 by choosing every one peak from four consecutive ones.

Although the data originally was from the physionet website, it was sampled at 125 Hz and stored as MATLAB files in the UCI repository [23]. The total dump file is divided into three separate MATLAB data files, each containing 3000 admission scenarios with ECG, PPG, and ABP data in separate rows. So, in this research, the first 1000 samples were considered to make it homogeneous. In Figure 11, the distribution of blood pressure over a total population of 12,000 is shown.



**Figure 11.** The range of blood pressure over all populations.

Also in Figure 12, an example of raw data representation from a patient of three vital signs is shown.



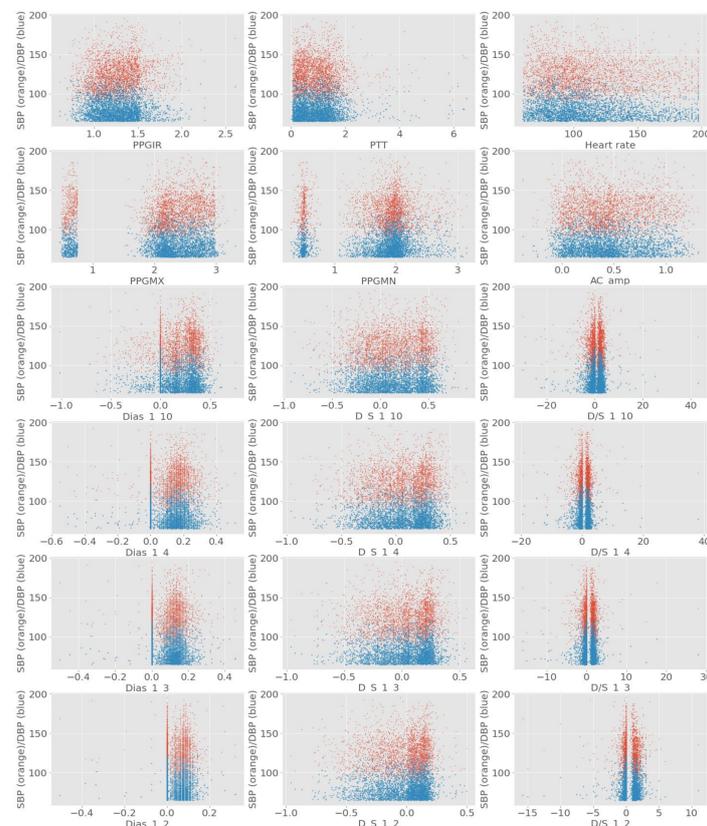
**Figure 12.** Example of a raw data from a patient with first 1000 samples.

### 3.2. Data Analysis

The raw signals from the dataset contain the three attributes but along with the signals, there were also some unwanted artifacts present.

Before removing unwanted noises, the dataset was manually checked for missing, outlier, or unusually different data points due to possible anomalies in the measurement system. From the 12,000-admission dataset of three attributes, the selected number of data was 3295. In the previously mentioned steps in the pre-processing subsection, the section was used to get rid of the noises and prepare the signal data ready for the data extraction phase. Since the FIR filter is linear but ECG is a non-linear signal, there is a possibility that there would be some delay involved as a result [130]. There were delays observed due to the use of the FIR filter. Using  $F_s$  and filter order, the delay comes as 0.196 s. If we had decided to increase the order of the filter, the time delay would have increased, which is a disadvantage of the FIR filter. Although this time delay did not hamper the calculation of heart rate or R peak detection but considering any other heart condition other than heart rate variability, we had to fix the delay. So, using the FIR filter, the delay does not distort the signal waveform, and it is possible to compensate for the delay simply by shifting the filtered signal by a specific sample to match the filtered version of the original signal waveform.

In the feature extraction phase, importance was given to the feature's biological significance regarding a change in blood pressure. The initial list of features to be extracted mostly consisted of different time intervals and amplitudes of specific wave sections of ECG and PPG waves. In Table 1, the detailed lists of features were given. The goal of the feature selection method was to use linearly and nonlinearly correlated features to be prioritized. Initially, the ranking using Pearson's linear correlation coefficient is shown in Table 2. In addition, in Figure 13 some of the distributions of the extracted features against their respective SBP and DBP values are depicted.



**Figure 13.** The distribution of SBP and DBP against some of the extracted features.

**Table 1.** The initial feature list from the feature extraction phase.

Name of Feature	Detail of Feature
HR	Rate of heartbeat per minute
PPGMX & PPGMN	Maximum—Minimum amplitude of the peak value of the PPG wave
AC_amp and PPGIR	The magnitude of the AC part and ratio of Maximum—Minimum value of PPG
PTT	The time gap between corresponding ECG and PPG signal
S_T and DT	Systolic and diastolic time
Dias, D_S, D/S of_1_10	DT with one-tenth—maximum value, DT +ST with one-tenth—maximum value, Ratio of DT- ST with one-tenth of maximum value
Dias, D_S, D/S of_1_4	DT with one-fourth—maximum value, DT +ST with one-fourth—maximum value, Ratio of DT- ST with one-fourth of maximum value
Dias, D_S, D/S of_1_3	DT with one-third—maximum value, DT +ST with one-third—maximum value k, Ratio of DT- ST with one-third of maximum value
Dias, D_S, D/S of_1_2	DT with half of maximum value, DT +ST with half of maximum value, Ratio of DT- ST with half of maximum value
D_S_2_3 and 3_4	DT + ST with 2/3 and 3/4 of maximum value
ECG time interval and amplitude	QRS, P, T wave intervals and peak amplitudes

**Table 2.** The ranking based on Pearson’s correlation coefficient.

SL	SBP	DBP
1	PPGIR	Heart rate
2	Dias_1_10	PPGMN
3	S_T	PPGIR
4	D_S_1_10	PPGMX
5	Dias_1_4	Dias_1_10
6	AC_amp	D_S_1_10
7	Dias_1_3	PTT
8	D_S_1_4	D/S_1_10
9	Dias_1_2	AC_amp
10	PPGMN	Dias_1_4
11	D_S_1_3	Dias_1_2
12	D_S_1_2	D_S_1_4

The linear correlation reduces →

The implementation of fast correlation using MSU and ReliefF-MSU algorithms was done using python language. The top 12 features were finalized based on the combined ranking of feature selection algorithms which were implemented based on the main rule of maximizing the relevance with the target variable but reducing the number of redundant features at the same time. It was found that using the proposed technique reducing the feature number to less than 12 produced results where the accuracy was less than the intended margin. In Figure 14, the heat map for the finalized feature for SBP and DBP is shown. Although the heat map only considers the linearly correlated coefficient and does not represent the correlation in terms of nonlinearly correlated features, it is evident from the heat map that even from the linear perspective the correlations among finalized features are quite small.

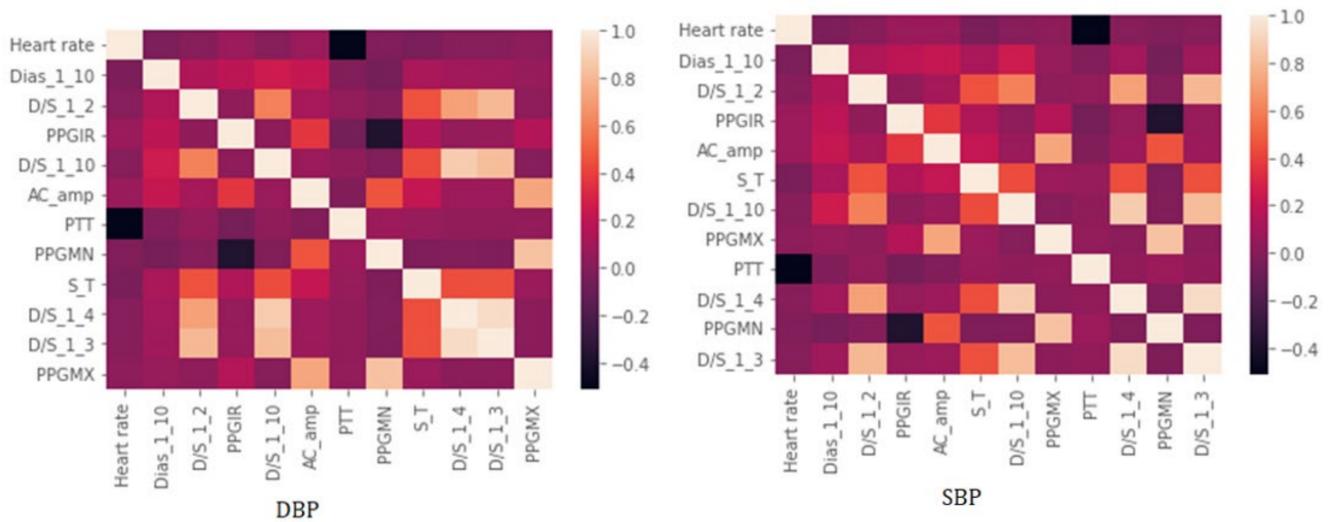


Figure 14. The Pearson’s coefficient value among finalized features using a heat map.

In this paper, the regularized techniques were applied because they keep the variance in control and bring a balance between bias–variance convex relationships. In addition, by nature regularization favors simpler models by addressing issues like variance-bias trade-off, multicollinearity, sparse data handling, and partly feature selection. Comparing popular models such as lasso and ridge where lasso assigns a penalty to reduce the number of variables and ridge assigns a penalty to reduce the impact of overemphasized variables, lasso showed some inconsistency due to the presence of multicollinearity even in smaller degree and ridge showed some inconsistency due to the lack of ability to reduce any variable to zero even when the situation required it. So, there must be a trade-off between the penalties coming from both lasso and ridge regression to make it balanced. The result of the processes which are coefficients or weights of estimates is shown in Figures 15 and 16.

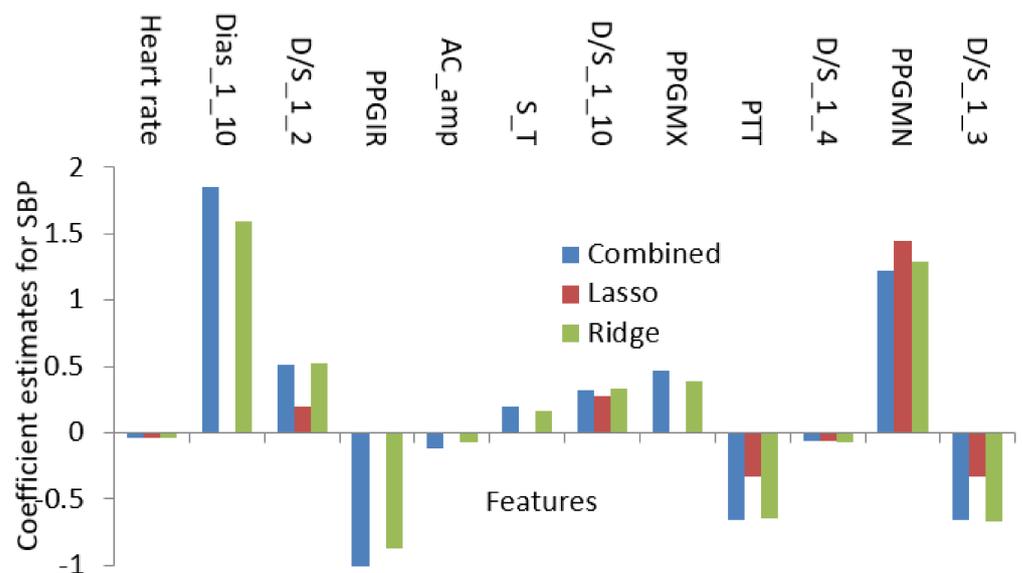


Figure 15. The value of coefficient for penalty-based regression for SBP.

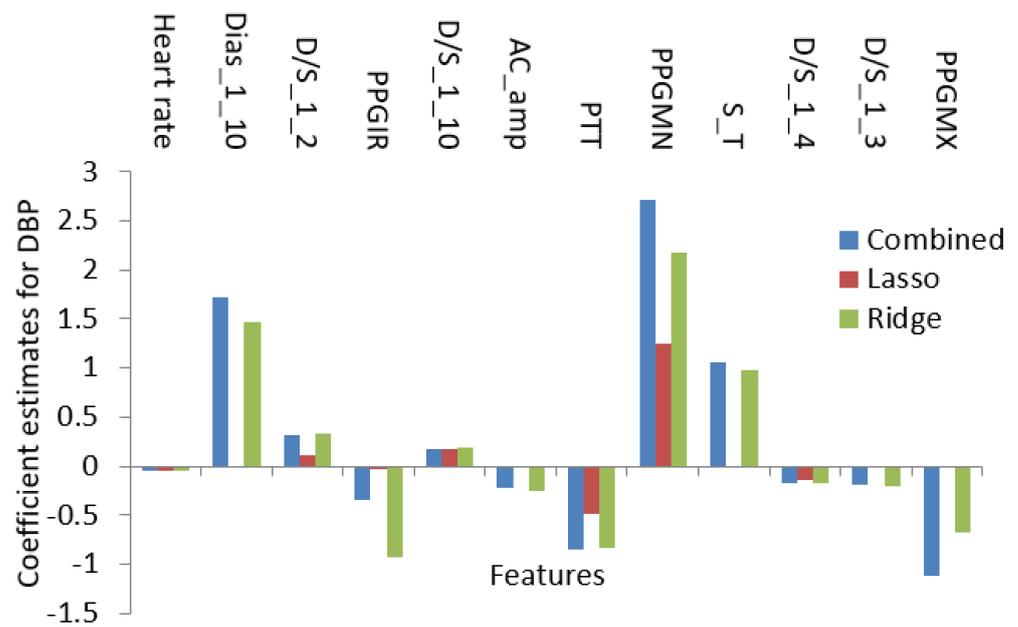


Figure 16. The value of the coefficient for penalty-based regression for DBP.

Using the features in Table 1 as the initial set, the optimization algorithms produced a result with the features listed in Table 2. The features in Table 2 encompass both linear and nonlinear relationships and using only those features the regression analysis was conducted. So, Figures 15 and 16 contain the features which not only have shown linear dependence but also from the aspects of both linear and nonlinear relationships.

In Figure 17, the coefficient values for features from the proposed method and Pearson’s method are shown. Few features were found common in both methods. However, more than half of the numbers of features are different. This variation in the selection of features proves the fact that using the linear relationship will only result in missing some more relevant features. For example, although in many studies the linear relationship between pulse transit time (PTT) and blood pressure was assumed, recently it has been proven that they are nonlinearly related [131]. In this study for SBP, the coefficient value is zero for the PPT feature when we use Pearson’s method. However, in the proposed method, the PTT coefficient is relatively significant confirming the findings. Similar conclusions can be drawn regarding DBP, as shown in Figure 18.

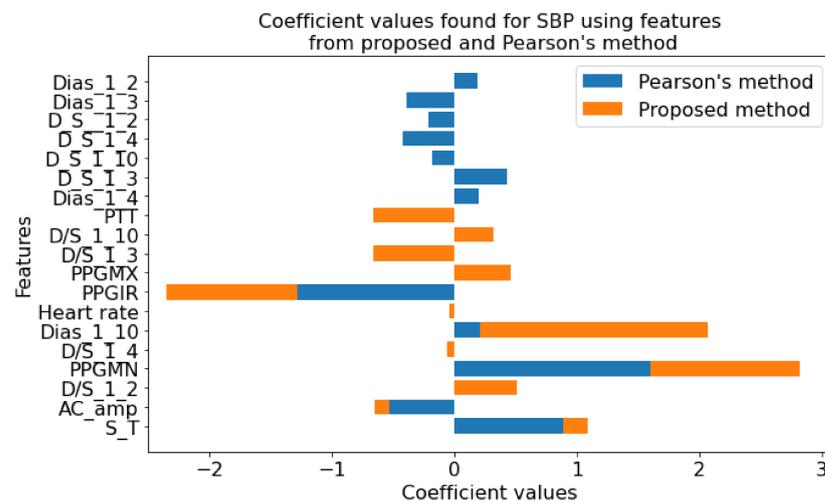
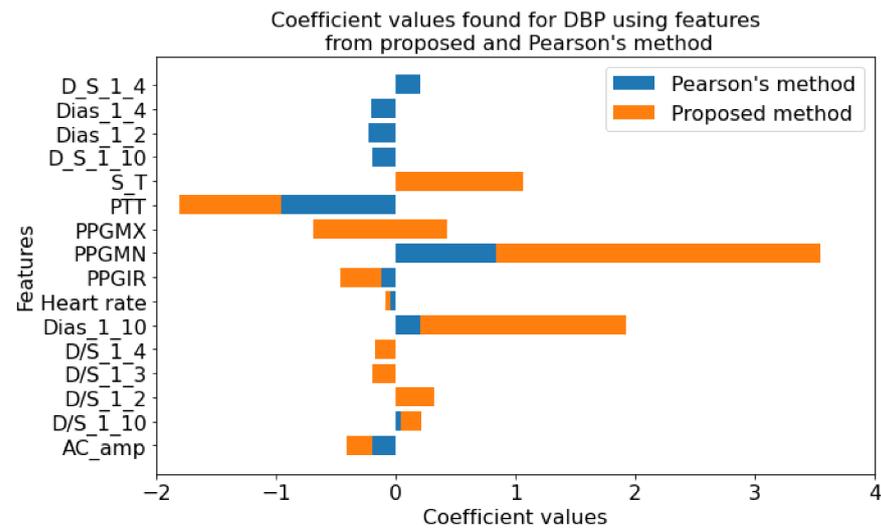


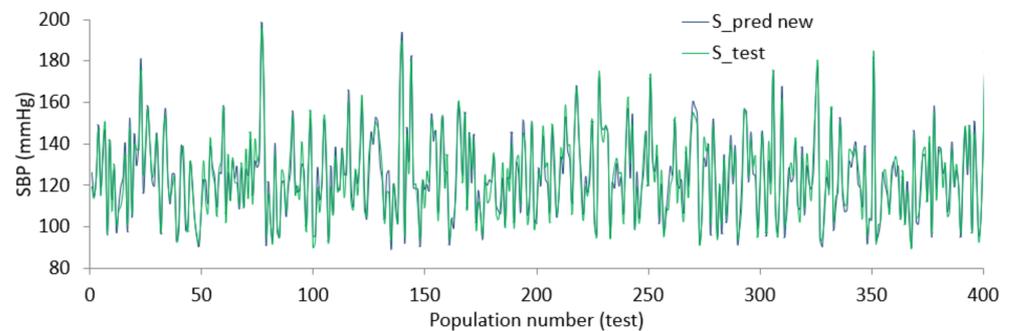
Figure 17. Coefficient values found for SBP using features from proposed method and Pearson’s method.



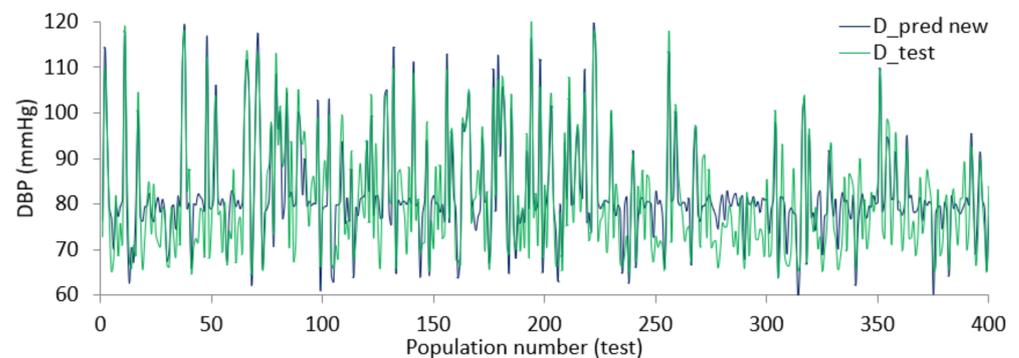
**Figure 18.** Coefficient values found for DBP using features from proposed method and Pearson's method.

#### 4. Result and Performance Evaluation

Using the coefficient values from the combination of ridge and lasso regression, the final relation between finalized features and SBP or DBP is defined. The regularized BP models were tested using data from the dataset resulting in the diagrams shown in Figures 19 and 20. From the Figures, there is evidence that the predicted SBD and DBP closely follow the test value in most cases.



**Figure 19.** Predicted value against actual value for SBP.



**Figure 20.** Predicted value against actual value for DBP.

The error varies within (+/-) 15 mmHg in a worst-case scenario, for more than half of the test points the error remains within (+/-) 5 mmHg. The size of the dataset was large and although some parts of the dataset were ignored due to lack of records or inconsistent

values, the performance shown by regularized regression using an optimized feature list was still promising. Since each of the feature selection algorithms works from different aspects, the outcome of the algorithms provides an optimized feature list. In addition, since the objective of this study is to bring forth a generalized relation and not a person-specific one, it was expected that there would be some discrepancies between the predicted value and test value.

The summary of the study in terms of root mean square error (RMSE) and mean absolute error (MAE) and the comparison with similar studies, where feature optimization and regression relations were used, are depicted in Table 3. Using the proposed technique, the performance indices MAE and RMSE improved significantly compared to the Pearson-based method. This performance improvement is the result of the inclusion of both linear and nonlinear relationships in the feature selection algorithm and the use of a penalty-based regression algorithm. The MAE and RSME for SBP were better than those of DBP for the proposed method. The main reason behind this difference is that nonlinearity of peripheral arterial compliance affects the dependency of PTT and SBP more than DBP [132]. In addition, by measuring hemodynamic parameters such as stroke volume and total peripheral resistance another study proved that the linear correlation is higher for DBP compared to SBP with pulse wave velocity and arterial stiffness parameters [133]. So, since the proposed method took linear and nonlinear relationships into account, the amount of improvement happened more for SBP than DBP. The number of features replaced from Pearson’s method to the proposed method also confirmed that for SBP there was a higher number of features needed to be replaced than DBP.

The value of MAE from the proposed method resulted slightly higher than the values from studies [134,135]. The experimental model of [136] was prepared using factors such as viscosity of arterial wall and pulse wave reflection but their number of subjects and age range were extremely small. They have used only 16 subjects with an age range of 19 to 25 only which lessened the probability of variation in both SBP and DBP. Replication of their technique was done using data used in this paper and the MAE was found as 9.8 mmHg for SBP and 9.3 mmHg for DBP. This experiment proves the necessity of a dataset with large variation in participants’ age and many records. The Framingham heart study [137], which followed blood pressure for more than 30 years, agreed that although SBP continuously increases as age goes from 30 to 84 and over, DBP slowly decreases after 50 and onward. In the study [138], out of all the participants, 17 were hypertensive and 12 were hypotensive. According to the Framingham study, the selection of age group from the participant will result in lower variation in DBP compared to SBP which explains the small value of MAE in DBP for [138]. Compared to those, the dataset used in this paper consisted of a much wider age range (20–100) as well as a higher number of patients [134]. In addition, in [138] the training and testing include measurement from the same 85 participants, which means the model has already “seen” those data before and might result in significant overfitting.

**Table 3.** Comparing the performance from different contemporary methods of BP measurement.

Method	MAE (mmHg)		Year
	SBP	DBP	
Using PTT [23]	6.85	6.35	2015
Peripheral PTT [136]	6.72	4.53	2016
Heart rate, pulse arrival time [135]	11.6	5.34	2017
PPT and IPG sensor [139]	8.55	5.07	2018
Feature from photoplethysmography wave [140]	9.42	6.87	2019
Using gradient boosting 1. and graph-theoretic algorithms [141]	8.09	5.49	2020
Regression using Pearson’s coefficients	9.50	10.10	2021
Proposed technique	4.27 (RMSE 5.28)	5.01 (RMSE 5.98)	2022

In addition, another significant criterion that was maintained while populating Table 4 is techniques or studies which used a large dataset with sufficient variations in data rather than the use of the selective and smaller dataset. In addition, using grading criteria proposed by BHS, the cumulative percentage and mean absolute error were measured with intervals as 5, 10, 15 mmHg. The minimum standard achieved by our proposed technique is grade B as shown in Table 4.

**Table 4.** Proposed method's performance according to BHS.

		BHS		
		CP—5 mmHg	CP—10 mmHg	CP—15 mmHg
BHS	Grade A	60%	85%	95%
	Grade B	50%	75%	90%
	Grade C	40%	65%	85%
Result from this study	SBP	65.29%	91.74%	99.87%
	DBP	58.37%	89.11%	99.63%

## 5. Conclusions and Discussion

The objective of this study was to propose and analyze the performance of a novel blood pressure measurement technique using an MSU-based optimized feature selection method combined with penalty-based regression. Researchers have been studying all relevant biomedical signals, such as ECG, PPG, PCG, BCG, etc., to extract necessary information and trying to come up with relations with change in blood pressure, but acceptable accuracy and consistency remain difficult to achieve. This paper proposed to improve accuracy and consistency by using a hybrid feature selection algorithm based on MSU so only relevant but nonredundant features from biomedical signals were used, this would make the input for the learning system optimized and reduce the chance of overfitting. The whole process initiates with biologically significant features from ECG and PPG signals which were extracted after necessary pre-processing steps. Using the advantage of both nonlinear and linear relations from multivariate symmetric uncertainty with the fast correlation and ReliefF algorithms, the optimized feature list was generated. Later, the finalized feature list was used for a penalty-based regularized regression to produce a general relation between the features and systolic/diastolic blood pressure. The regularized regression provides a balance between bias and variance to optimize the error or difference between the actual and predicted result.

In this study, the features were chosen based on their morphological features and their respective significance. To include statistical or spectral features there must be sufficient available information regarding how those features impact the change in blood pressure. Using the MIMIC II dataset, the regression model was trained and tested. The results showed improved performance in terms of RMSE and MAE compared to contemporary studies.

One significant limitation of this study is the assumption related to the arterial stiffness from the PPG wave pattern. Although this is not the ideal way, the assumptions were made to consider the effect the arterial stiffness without using additional sensors or invasive methods to measure arterial dimensions. Due to the assumptions, measurement accuracy was affected to some extent. The selection of features was done to compensate for that effect by choosing features specifically related to arterial stiffness. So, compared to other studies the effect of the assumptions was much lower in this study. It is a trade-off between assumptions and invasive sensors to get precise arterial information. Personal calibration with sufficient data should be an alternative solution for this trade-off.

Future work will use the developed measurement model and tune it using patient data for further calibration.

**Author Contributions:** Conceptualization, M.M.R.K.M.; Investigation, M.M.R.K.M.; Methodology, M.M.R.K.M.; Project administration, A.T.A.; Supervision, A.T.A.; Visualization, M.M.R.K.M.; Writing—original draft, M.M.R.K.M.; Writing—review and editing, A.T.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external or internal funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Publicly archived datasets analyzed which can be found in the following website: <https://archive.physionet.org/mimic2/> (accessed on 15 January 2021).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Kearney, P.M.; Whelton, M.; Reynolds, K.; Muntner, P.; Whelton, P.K.; He, J. Global burden of hypertension: Analysis of worldwide data. *Lancet* **2005**, *365*, 217–223. [[CrossRef](#)]
2. World Health Organization. *World Health Statistics 2016: Monitoring Health for the SDGs Sustainable Development Goals*; World Health Organization: Geneva, Switzerland, 2016.
3. Burt, V.L.; Whelton, P.; Roccella, E.J.; Brown, C.; Cutler, J.A.; Higgins, M.; Horan, M.J.; Labarthe, D. Prevalence of hypertension in the US adult population: Results from the Third National Health and Nutrition Examination Survey, 1988–1991. *Hypertension* **1995**, *25*, 305–313. [[CrossRef](#)] [[PubMed](#)]
4. Irigoyen, M.-C.; DE Angelis, K.; Dos Santos, F.; Dartora, D.R.; Rodrigues, B.; Consolim-Colombo, F.M. Hypertension, Blood Pressure Variability, and Target Organ Lesion. *Curr. Hypertens. Rep.* **2016**, *18*, 31. [[CrossRef](#)] [[PubMed](#)]
5. Tientcheu, D.D.; Ayers, C.; Das, S.; McGuire, D.; De Lemos, J.; Khera, A.; Victor, R.; Kaplan, N.; Vongpatanasin, W. Target organ complications and cardiovascular events associated with masked hypertension and white coat hypertension: Analysis from the Dallas heart study. *J. Am. Soc. Hypertens.* **2015**, *9*, e1. [[CrossRef](#)]
6. Drzewiecki, G.; Hood, R.; Apple, H. Theory of the oscillometric maximum and the systolic and diastolic detection ratios. *Ann. Biomed. Eng.* **1994**, *22*, 88–96. [[CrossRef](#)]
7. Perloff, D.; Grim, C.; Flack, J.; Frohlich, E.D.; Hill, M.; McDonald, M.; Morgenstern, B.Z. Human blood pressure determination by sphygmomanometry. *Circulation* **1993**, *88*, 2460–2470. [[CrossRef](#)]
8. Wang, G.; Atef, M.; Lian, Y. Towards a Continuous Non-Invasive Cuffless Blood Pressure Monitoring System Using PPG: Systems and Circuits Review. *IEEE Circuits Syst. Mag.* **2018**, *18*, 6–26. [[CrossRef](#)]
9. McGhee, B.H.; Bridges, E.J. Monitoring arterial blood pressure: What you may not know. *Crit. Care Nurse* **2002**, *22*, 60–79. [[CrossRef](#)]
10. Chan, G.; Cooper, R.; Hosanee, M.; Welykholowa, K.; Kyriacou, P.A.; Zheng, D.; Allen, J.; Abbott, D.; Lovell, N.H.; Fletcher, R.; et al. Multi-Site Photoplethysmography Technology for Blood Pressure Assessment: Challenges and Recommendations. *J. Clin. Med.* **2019**, *8*, 1827. [[CrossRef](#)]
11. Chen, W.; Kobayashi, T.; Ichikawa, S.; Takeuchi, Y.; Togawa, T. Continuous estimation of systolic blood pressure using the pulse arrival time and intermittent calibration. *Med. Biol. Eng. Comput.* **2000**, *38*, 569–574. [[CrossRef](#)]
12. Poon, C.; Zhang, Y. Cuff-less and Noninvasive Measurements of Arterial Blood Pressure by Pulse Transit Time. In Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Berlin, Germany, 23–27 July 2005; Volume 2005, pp. 5877–5880.
13. Kumar, N.; Agrawal, A.; Deb, S. Cuffless BP measurement using a correlation study of pulse transient time and heart rate. In Proceedings of the 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Delhi, India, 24–27 September 2014; pp. 1538–1541.
14. Chen, Y.; Wen, C.; Tao, G.; Bi, M.; Li, G. Continuous and Noninvasive Blood Pressure Measurement: A Novel Modeling Methodology of the Relationship between Blood Pressure and Pulse Wave Velocity. *Ann. Biomed. Eng.* **2009**, *37*, 2222–2233. [[CrossRef](#)] [[PubMed](#)]
15. Chen, Y.; Wen, C.; Tao, G.; Bi, M. Continuous and Noninvasive Measurement of Systolic and Diastolic Blood Pressure by One Mathematical Model with the Same Model Parameters and Two Separate Pulse Wave Velocities. *Ann. Biomed. Eng.* **2011**, *40*, 871–882. [[CrossRef](#)] [[PubMed](#)]
16. Mamun, M.M.R.K. Cuff-less blood pressure measurement based on hybrid feature selection algorithm and multi-penalty regularized regression technique. *Biomed. Phys. Eng. Express* **2021**, *7*, 065030. [[CrossRef](#)] [[PubMed](#)]
17. Mamun, M.M.R.K.; Alouani, A. Using photoplethysmography & ECG towards a non-invasive cuff less blood pressure measurement technique. In Proceedings of the 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), Edmonton, AB, Canada, 5–8 May 2019; pp. 1–4.

18. Forouzanfar, M.; Ahmad, S.; Batkin, I.; Dajani, H.R.; Groza, V.Z.; Bolic, M. Model-Based Mean Arterial Pressure Estimation Using Simultaneous Electrocardiogram and Oscillometric Blood Pressure Measurements. *IEEE Trans. Instrum. Meas.* **2015**, *64*, 2443–2452. [[CrossRef](#)]
19. Ma, T.; Zhang, Y. A Correlation Study on the Variabilities in Pulse Transit Time, Blood Pressure, and Heart Rate Recorded Simultaneously from Healthy Subjects. In Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, 17–18 January 2005; Volume 2006, pp. 996–999.
20. Wagenseil, J.E.; Mecham, R.P. Elastin in Large Artery Stiffness and Hypertension. *J. Cardiovasc. Transl. Res.* **2012**, *5*, 264–273. [[CrossRef](#)]
21. Sharma, M.; Barbosa, K.; Ho, V.; Griggs, D.; Ghirmai, T.; Krishnan, S.K.; Hsiai, T.K.; Chiao, J.-C.; Cao, H. Cuff-Less and Continuous Blood Pressure Monitoring: A Methodological Review. *Technologies* **2017**, *5*, 21. [[CrossRef](#)]
22. McCombie, D.B.; Reisner, A.T.; Asada, H.H. Adaptive blood pressure estimation from wearable PPG sensors using peripheral artery pulse wave velocity measurements and multi-channel blind identification of local arterial dynamics. In Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, USA, 31 August–3 September 2006; pp. 3521–3524.
23. Kachuee, M.; Kiani, M.M.; Mohammadzade, H.; Shabany, M. Cuff-less high-accuracy calibration-free blood pressure estimation using pulse transit time. In Proceedings of the 2015 IEEE International Symposium on Circuits and Systems (ISCAS), Lisbon, Portugal, 24–27 May 2015; pp. 1006–1009.
24. Esmaelpoor, J.; Moradi, M.; Kadkhodamohammadi, A. Cuffless blood pressure estimation methods: Physiological model parameters versus machine-learned features. *Phys. Meas.* **2021**, *42*, 035006. [[CrossRef](#)]
25. Yoon, Y.-Z.; Kang, J.M.; Kwon, Y.; Park, S.; Noh, S.; Kim, Y.; Park, J.; Hwang, S.W. Cuff-Less Blood Pressure Estimation Using Pulse Waveform Analysis and Pulse Arrival Time. *IEEE J. Biomed. Health Inform.* **2017**, *22*, 1068–1074. [[CrossRef](#)]
26. Xu, Z.; Liu, J.; Chen, X.; Wang, Y.; Zhao, Z. Continuous blood pressure estimation based on multiple parameters from electrocardiogram and photoplethysmogram by Back-propagation neural network. *Comput. Ind.* **2017**, *89*, 50–59. [[CrossRef](#)]
27. Tanveer, M.S.; Hasan, M.K. Cuffless blood pressure estimation from electrocardiogram and photoplethysmogram using waveform based ANN-LSTM network. *Biomed. Signal Process. Control* **2019**, *51*, 382–392. [[CrossRef](#)]
28. Chang, C.; Lin, C. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol. (TIST)* **2011**, *2*, 1–27. [[CrossRef](#)]
29. Liu, M.; Po, L.-M.; Fu, H. Cuffless Blood Pressure Estimation Based on Photoplethysmography Signal and Its Second Derivative. *Int. J. Comput. Theory Eng.* **2017**, *9*, 202–206. [[CrossRef](#)]
30. Chao, P.; Wu, C.; Nguyen, D.; Nguyen, B.; Huang, P.; Le, V. The Machine Learnings Leading the Cuffless PPG Blood Pressure Sensors Into the Next Stage. *IEEE Sens. J.* **2021**, *21*, 12498–12510.
31. Safar, M.E.; Asmar, R.; Benetos, A.; Blacher, J.; Boutouyrie, P.; Lacolley, P.; Laurent, S.; London, G.; Pannier, B.; Protogerou, A.; et al. Interaction between hypertension and arterial stiffness: An expert reappraisal. *Hypertension* **2018**, *72*, 796–805. [[CrossRef](#)] [[PubMed](#)]
32. Shaltis, P.; Reisner, A.; Asada, H. Calibration of the Photoplethysmogram to Arterial Blood Pressure: Capabilities and Limitations for Continuous Pressure Monitoring. In Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 30 August–3 September 2005; Volume 2005, pp. 3970–3973.
33. Elgendi, M. On the Analysis of Fingertip Photoplethysmogram Signals. *Curr. Cardiol. Rev.* **2012**, *8*, 14–25. [[CrossRef](#)] [[PubMed](#)]
34. Yoon, G.; Lee, J.Y.; Jeon, K.J.; Park, K.K.; Yeo, H.S.; Hwang, H.T.; Kim, H.S.; Hwang, I.D. Multiple diagnosis based on photoplethysmography: Hematocrit, SpO<sub>2</sub>, pulse, and respiration. In *Optics in Health Care and Biomedical Optics: Diagnostics and Treatment*; International Society for Optics and Photonics: Bellingham, WA USA, 2002; Volume 4916, pp. 185–188.
35. Reyes, I.; Nazeran, H.; Franco, M.; Haltiwanger, E. Wireless photoplethysmographic device for heart rate variability signal acquisition and analysis. In Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012; Volume 2012, pp. 2092–2095.
36. Chen, S.; Ji, Z.; Wu, H.; Xu, Y. A Non-Invasive Continuous Blood Pressure Estimation Approach Based on Machine Learning. *Sensors* **2019**, *19*, 2585. [[CrossRef](#)]
37. Ripoll, V.R.; Vellido, A. Blood Pressure Assessment with Differential Pulse Transit Time and Deep Learning: A Proof of Concept. *Kidney Dis.* **2018**, *5*, 23–27. [[CrossRef](#)] [[PubMed](#)]
38. El-Hajj, C.; Kyriacou, P. Cuffless blood pressure estimation from PPG signals and its derivatives using deep learning models. *BSPC* **2021**, *70*, 102984. [[CrossRef](#)]
39. Shimazaki, S.; Bhuiyan, S.; Kawanaka, H.; Oguri, K. Features Extraction for Cuffless Blood Pressure Estimation by Autoencoder from Photoplethysmography. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 17–21 July 2018; Volume 2018, pp. 2857–2860.
40. Wang, L.; Zhou, W.; Xing, Y.; Zhou, X. A Novel Neural Network Model for Blood Pressure Estimation Using Photoplethysmography without Electrocardiogram. *J. Healthc. Eng.* **2018**, *2018*, 1–9. [[CrossRef](#)]
41. Miao, F.; Fu, N.; Zhang, Y.; Ding, X.-R.; Hong, X.; He, Q.; Li, Y. A Novel Continuous Blood Pressure Estimation Approach Based on Data Mining Techniques. *IEEE J. Biomed. Health Inform.* **2017**, *21*, 1730–1740. [[CrossRef](#)]
42. Suzuki, S.; Oguri, K. Cuffless blood pressure estimation by error-correcting output coding method based on an aggregation of AdaBoost with a photoplethysmograph sensor. In Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Guadalajara, Mexico, 3–6 September 2009; Volume 2009, pp. 6765–6768.

43. Teng, X.F.; Zhang, Y.T. Continuous and noninvasive estimation of arterial blood pressure using a photoplethysmographic approach. In Proceedings of the IEEE Engineering in Medicine and Biology Society Conference, Cancun, Mexico, 17–21 September 2003; pp. 3153–3156.
44. Ruiz-Rodríguez, J.C.; Ruiz-Sanmartín, A.; Ribas, V.; Caballero, J.; García-Roche, A.; Riera, J.; Nuvials, X.; De Nadal, M.; De Solà-Morales, O.; Serra, J.; et al. Innovative continuous non-invasive cuffless blood pressure monitoring based on photoplethysmography technology. *Intensive Care Med.* **2013**, *39*, 1618–1625. [[CrossRef](#)]
45. Kurylyak, Y.; Lamonaca, F.; Grimaldi, D. A Neural Network-based method for continuous blood pressure estimation from a PPG signal. In Proceedings of the 2013 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Minneapolis, MN, USA, 6–9 May 2013; pp. 280–283.
46. ISO 81060-2/ANSI-AAMI; Non-Invasive Sphygmomanometers—Part 2: Clinical Investigation of Automated Measurement Type ANSI/AAMI. Association for the Advancement of Medical Instrumentation: Arlington, VA, USA, 2013.
47. Jiao, Z.; Gao, X.; Wang, Y.; Li, J.; Xu, H. Deep Convolutional Neural Networks for mental load classification based on EEG data. *Pattern Recognit.* **2018**, *76*, 582–595. [[CrossRef](#)]
48. Antipov, G.; Baccouche, M.; Berrani, S.-A.; Dugelay, J.-L. Effective training of convolutional neural networks for face-based gender and age prediction. *Pattern Recognit.* **2017**, *72*, 15–26. [[CrossRef](#)]
49. Zhang, Q.; Yang, L.T.; Chen, Z.; Li, P. A survey on deep learning for big data. *Inf. Fusion* **2018**, *42*, 146–157. [[CrossRef](#)]
50. Mamun, M.M.R.K. Assessment of Potential Primary and Recurrent Ischemic Stroke by Detecting Atrial Fibrillation using 1D-CNN and CHA2DS2-VA Score. In Proceedings of the 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Houston, TX, USA, 9–12 December 2021; pp. 1963–1968.
51. Mamun, M.M.R.K.; Alouani, A. Arrhythmia Classification Using Hybrid Feature Selection Approach and Ensemble Learning Technique. In Proceedings of the 2021 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Online, 12–17 September 2021; pp. 1–6.
52. Mamun, M.M.R.K.; Alouani, A. Diagnosis of STEMI and Non-STEMI Heart Attack using Nature-inspired Swarm Intelligence and Deep Learning Techniques. *J. Biomed. Eng. Bios. (JBEB)* **2020**, *6*, 1–8.
53. Mamun, M.M.K.; Alouani, A. FA-1D-CNN implementation to improve diagnosis of heart disease risk level. In Proceedings of the 6th World Congress on Engineering and Computer Systems and Sciences, San Francisco, CA, USA, 25–27 October 2020; pp. 122.1–122.9.
54. Liang, Y.; Chen, Z.; Ward, R.; Elgendi, M. Photoplethysmography and Deep Learning: Enhancing Hypertension Risk Stratification. *Biosensors* **2018**, *8*, 101. [[CrossRef](#)]
55. Su, P.; Ding, X.-R.; Zhang, Y.-T.; Liu, J.; Miao, F.; Zhao, N. Long-term blood pressure prediction with deep recurrent neural networks. In Proceedings of the 2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), Las Vegas, NE, USA, 4–7 March 2018; pp. 323–328.
56. Shimazaki, S.; Kawanaka, H.; Ishikawa, H.; Inoue, K.; Oguri, K. Cuffless Blood Pressure Estimation from only the Waveform of Photoplethysmography using CNN. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; Volume 2019, pp. 5042–5045.
57. Qiu, Y.; Liu, D.; Yang, G.; Qi, D.; Lu, Y.; He, Q.; Qian, X.; Li, X.; Cao, Y.; Shuai, J. Cuffless blood pressure estimation based on composite neural network and graphics information. *Biomed. Signal Process. Control* **2021**, *70*, 103001. [[CrossRef](#)]
58. Athaya, T.; Choi, S. An Estimation Method of Continuous Non-Invasive Arterial Blood Pressure Waveform Using Photoplethysmography: A U-Net Architecture-Based Approach. *Sensors* **2021**, *21*, 1867. [[CrossRef](#)]
59. Aguirre, N.; Grall-Maës, E.; Cymberknop, L.; Armentano, R. Blood Pressure Morphology Assessment from Photoplethysmogram and Demographic Information Using Deep Learning with Attention Mechanism. *Sensors* **2021**, *21*, 2167. [[CrossRef](#)]
60. El-Hajj, C.; Kyriacou, P. A review of machine learning techniques in photoplethysmography for the non-invasive cuff-less measurement of blood pressure. *Biomed. Signal Process. Control* **2020**, *58*, 101870. [[CrossRef](#)]
61. Schrupf, F.; Frenzel, P.; Aust, C.; Osterhoff, G.; Fuchs, M. Assessment of Non-Invasive Blood Pressure Prediction from PPG and rPPG Signals Using Deep Learning. *Sensors* **2021**, *21*. [[CrossRef](#)]
62. Saleh, H.; Younis, E.M.G.; Sahal, R.; Ali, A.A. Predicting Systolic Blood Pressure in Real-Time Using Streaming Data and Deep Learning. *Mob. Netw. Appl.* **2021**, *26*, 326–335. [[CrossRef](#)]
63. Li, Y.-H.; Harfiya, L.N.; Purwandari, K.; Lin, Y.-D. Real-Time Cuffless Continuous Blood Pressure Estimation Using Deep Learning Model. *Sensors* **2020**, *20*, 5606. [[CrossRef](#)] [[PubMed](#)]
64. Eom, H.; Lee, D.; Han, S.; Hariyani, Y.S.; Lim, Y.; Sohn, I.; Park, K.; Park, C. End-To-End Deep Learning Architecture for Continuous Blood Pressure Estimation Using Attention Mechanism. *Sensors* **2020**, *20*, 2338. [[CrossRef](#)] [[PubMed](#)]
65. Miao, F.; Wen, B.; Hu, Z.; Fortino, G.; Wang, X.-P.; Liu, Z.-D.; Tang, M.; Li, Y. Continuous blood pressure measurement from one-channel electrocardiogram signal using deep-learning techniques. *Artif. Intell. Med.* **2020**, *108*, 101919. [[CrossRef](#)]
66. Kasieczka, G.; Shih, D. Robust Jet Classifiers through Distance Correlation. *Phys. Rev. Lett.* **2020**, *125*, 122001. [[CrossRef](#)]
67. Hayashi, K.; Yamamoto, T.; Takahara, A.; Shirai, K. Clinical assessment of arterial stiffness with cardio-ankle vascular index: Theory and applications. *J. Hypertens.* **2015**, *33*, 1742–1757. [[CrossRef](#)] [[PubMed](#)]
68. Liberson, A.S.; Lillie, J.S.; Day, S.W.; Borkholder, D.A. A physics based approach to the pulse wave velocity prediction in compliant arterial segments. *J. Biomech.* **2016**, *49*, 3460–3466. [[CrossRef](#)] [[PubMed](#)]
69. Shiogai, Y.; Stefanovska, A.; McClintock, P. Nonlinear dynamics of cardiovascular ageing. *Phys. Rep.* **2010**, *488*, 51–110. [[CrossRef](#)]

70. Bolón-Canedo, V.; Alonso-Betanzos, A. Ensembles for feature selection: A review and future trends. *Inf. Fusion* **2019**, *52*, 1–12.
71. Krakovska, O.; Christie, G.; Sixsmith, A.; Ester, M.; Moreno, S. Performance comparison of linear and non-linear feature selection methods for the analysis of large survey datasets. *PLoS ONE* **2019**, *14*, e0213584. [[CrossRef](#)]
72. Loperfido, N. Canonical Correlations and Nonlinear Dependencies. *Symmetry* **2021**, *13*, 1308. [[CrossRef](#)]
73. Hartmann, V.; Liu, H.; Chen, F.; Hong, W.; Hughes, S.; Zheng, D. Toward Accurate Extraction of Respiratory Frequency from the Photoplethysmogram: Effect of Measurement Site. *Front. Physiol.* **2019**, *10*, 732. [[CrossRef](#)]
74. Bussy, C.; Boutouyrie, P.; Lacolley, P.; Challande, P.; Laurent, S. Intrinsic Stiffness of the Carotid Arterial Wall Material in Essential Hypertensives. *Hypertension* **2000**, *35*, 1049–1054. [[CrossRef](#)]
75. Huvneers, S.; Daemen, M.J.; Hordijk, P.L. Between Rho (k) and a hard place: The relation between vessel wall stiffness, endothelial contractility, and cardiovascular disease. *Circ. Res.* **2015**, *116*, 895–908. [[CrossRef](#)]
76. Laurent, S.; Briet, M.; Boutouyrie, P. Large and Small Artery Cross-Talk and Recent Morbidity-Mortality Trials in Hypertension. *Hypertension* **2009**, *54*, 388–392. [[CrossRef](#)]
77. Klabunde, R. *Cardiovascular Physiology Concepts*; Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2011.
78. Choi, W.; Cho, J. Correlation of peak time shift in blood pressure waveform and PPG based on compliance change analysis in RLC Windkessel model. *Curr. Opt. Photonics* **2017**, *1*, 529–537.
79. Elhaj, F.A.; Salim, N.; Harris, A.R.; Swee, T.T.; Ahmed, T. Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals. *Comput. Methods Programs Biomed.* **2016**, *127*, 52–63. [[CrossRef](#)]
80. de Chazal, P.; O'Dwyer, M.; Reilly, R. Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features. *IEEE Trans. Biomed. Eng.* **2004**, *51*, 1196–1206. [[CrossRef](#)]
81. Kher, R. Signal processing techniques for removing noise from ECG signals. *J. Biomed. Eng. Res.* **2019**, *3*, 1–9.
82. Watford, C. *Understanding ECG Filtering*; EMS 12-Lead: Hilton Head Island, SC, USA, 2014.
83. Pan, J.; Tompkins, W.J. A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.* **1985**, *3*, 230–236. [[CrossRef](#)]
84. Karavaev, A.S.; Borovik, A.S.; Borovkova, E.I.; Orlova, E.A.; Simonyan, M.A.; Ponomarenko, V.I.; Skazkina, V.V.; Gridnev, V.I.; Bezruchko, B.P.; Prokhorov, M.D.; et al. Low-frequency component of photoplethysmogram reflects the autonomic control of blood pressure. *Biophys. J.* **2021**. [[CrossRef](#)]
85. Alty, S.R.; Angarita-Jaimes, N.; Millasseau, S.C.; Chowienzyk, P. Predicting Arterial Stiffness from the Digital Volume Pulse Waveform. *IEEE Trans. Biomed. Eng.* **2007**, *54*, 2268–2275. [[CrossRef](#)]
86. Takazawa, K.; Tanaka, N.; Fujita, M.; Matsuoka, O.; Saiki, T.; Aikawa, M.; Tamura, S.; Ibukiyama, C. Assessment of Vasoactive Agents and Vascular Aging by the Second Derivative of Photoplethysmogram Waveform. *Hypertension* **1998**, *32*, 365–370. [[CrossRef](#)]
87. Aksoy, S.; Gürkan, U.; Öz, D.; Dayı, Ş.Ü.; Demirci, D.; Eksik, A.; Sayar, N.; Agirbasli, M. The Effects of Blood Pressure Lowering on P-Wave Dispersion in Patients with Hypertensive Crisis in Emergency Setting. *Clin. Exp. Hypertens.* **2010**, *32*, 486–489. [[CrossRef](#)]
88. Sun, G.; Zhou, Y.; Ye, N.; Wu, S.; Sun, Y. Independent associations of blood pressure and body mass index with interatrial block: A cross-sectional study in general Chinese population. *BMJ Open* **2019**, *9*, e029463. [[CrossRef](#)]
89. Ferrucci, A.; Canichella, F.; Battistoni, A.; Palano, F.; Francia, P.; Ciavarella, G.M.; Volpe, M.; Tocci, G. A Novel Electrocardiographic T-Wave Measurement (Tp-Te Interval) as a Predictor of Heart Abnormalities in Hypertension: A New Opportunity for First-Line Electrocardiographic Evaluation. *J. Clin. Hypertens.* **2015**, *17*, 441–449. [[CrossRef](#)]
90. Hassing, G.J.; van der Wall, H.; van Westen, G.; Kemme, M.; Adiyaman, A.; Elvan, A.; Burggraaf, J.; Gal, P. Blood pressure-related electrocardiographic findings in healthy young individuals. *Blood Press.* **2019**, *29*, 113–122. [[CrossRef](#)]
91. Akintunde, A.A.; Ayodele, O.E.; Opadijo, O.G.; Oyedeji, A.T.; Familoni, O.B. QT Interval prolongation and dispersion: Epidemiology and clinical correlates in subjects with newly diagnosed systemic hypertension in Nigeria. *J. Cardiovasc. Dis. Res.* **2012**, *3*, 290–295. [[CrossRef](#)]
92. Anigbogu, C.N.; Isichei, C.V.; Ajuluchukwu, J.N. Blood pressure, heart rate, cardiovascular reflexes and electrocardiographic changes in some hypertensive Nigerians. *Niger. J. Physiol. Sci.* **2012**, *27*, 23–27.
93. Anttila, I.; Nikus, K.; Nieminen, T.; Jula, A.; Reunanen, A.; Salomaa, V.; Kattainen, A.; Nieminen, M.S.; Lehtimäki, T.; Virtanen, V.; et al. Prevalence and prognostic value of poor R-wave progression in standard resting electrocardiogram in a general adult population. The Health 2000 Survey. *Ann. Med.* **2010**, *42*, 135–142. [[CrossRef](#)]
94. Aeschbacher, S.; Metin, F.; Bossard, M.; Schoen, T.; von Rotz, M.; Mettler, H.; Abächerli, R.; Risch, M.; Risch, L.; Conen, D. Relationships of electrocardiographic parameters with ambulatory hypertension in young and healthy adults. *Int. J. Cardiol.* **2015**, *202*, 300–304. [[CrossRef](#)]
95. Bekar, L.; Kalçık, M.; Kilci, H.; Çelik, O.; Yetim, M.; Doğan, T.; Önalın, O. Presence of fragmented QRS may be associated with complex ventricular arrhythmias in patients with essential hypertension. *J. Electrocardiol.* **2019**, *55*, 20–25. [[CrossRef](#)]
96. Bird, K.; Chan, G.; Lu, H.; Greeff, H.; Allen, J.; Abbott, D.; Menon, C.; Lovell, N.H.; Howard, N.; Chan, W.-S.; et al. Assessment of Hypertension Using Clinical Electrocardiogram Features: A First-Ever Review. *Front. Med.* **2020**, *7*. [[CrossRef](#)]
97. Chua, E.C.-P.; Redmond, S.; McDarby, G.; Heneghan, C. Towards Using Photo-Plethysmogram Amplitude to Measure Blood Pressure During Sleep. *Ann. Biomed. Eng.* **2010**, *38*, 945–954. [[CrossRef](#)]
98. Awad, A.A.; Haddadin, A.S.; Tantawy, H.; Badr, T.M.; Stout, R.G.; Silverman, D.G.; Shelley, K.H. The relationship between the photoplethysmographic waveform and systemic vascular resistance. *Int. J. Clin. Monit. Comput.* **2007**, *21*, 365–372. [[CrossRef](#)]

99. Wang, L.; Pickwell-MacPherson, E.; Liang, Y.P.; Zhang, Y. Noninvasive cardiac output estimation using a novel photoplethysmogram index. In Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, USA, 3–6 September 2009; Volume 2009, pp. 1746–1749. [[CrossRef](#)]
100. Martínez, G.; Howard, N.; Abbott, D.; Lim, K.; Ward, R.; Elgendi, M. Can Photoplethysmography Replace Arterial Blood Pressure in the Assessment of Blood Pressure? *J. Clin. Med.* **2018**, *7*, 316. [[CrossRef](#)]
101. Bashir, D.; Montañez, G.D.; Sehra, S.; Segura, P.S.; Lauw, J. An Information-Theoretic Perspective on Overfitting and Underfitting. In Proceedings of the Australasian Joint Conference on Artificial Intelligence, Canberra, ACT, Australia, 29–30 November 2020; Lecture Notes in Computer Science; Springer International Publishing: Cham, Switzerland, 2020; pp. 347–358.
102. Chandrashekar, G.; Sahin, F. A survey on feature selection methods. *Comput. Electr. Eng.* **2014**, *40*, 16–28. [[CrossRef](#)]
103. Tabakhi, S.; Najafi, A.; Ranjbar, R.; Moradi, P. Gene selection for microarray data classification using a novel ant colony optimization. *Neurocomputing* **2015**, *168*, 1024–1036. [[CrossRef](#)]
104. Kononenko, I. Estimating Attributes: Analysis and Extensions of RELIEF. In Proceedings of the European Conference on Machine Learning, Catania, Italy, 6–8 April 1994; Springer: Berlin/Heidelberg, Germany, 1994; pp. 171–182.
105. Robnik-Šikonja, M.; Kononenko, I. Theoretical and Empirical Analysis of ReliefF and RReliefF. *Mach. Learn.* **2003**, *53*, 23–69. [[CrossRef](#)]
106. Sen, R.; Mandal, A.K.; Goswami, S.; Chakraborty, B. A Comparative Study of the Stability of Filter based Feature Selection Algorithms. In Proceedings of the 2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST), Morioka, Japan, 23–25 October 2019; pp. 1–6.
107. Al-Ani, A.; Deriche, M.; Chebil, J. A new mutual information based measure for feature selection. *Intell. Data Anal.* **2003**, *7*, 43–57. [[CrossRef](#)]
108. Gu, X.; Guo, J. A feature subset selection algorithm based on equal interval division and three-way interaction information. *Soft Comput.* **2021**, *25*, 1–11. [[CrossRef](#)]
109. Venkatesh, B.; Anuradha, J. A Review of Feature Selection and Its Methods. *Cybern. Inf. Technol.* **2019**, *19*, 3–26. [[CrossRef](#)]
110. Speed, T. A Correlation for the 21st Century. *Science* **2011**, *334*, 1502–1503. [[CrossRef](#)] [[PubMed](#)]
111. Hardin, J.; Mitani, A.; Hicks, L.; VanKoten, B. A robust measure of correlation between two genes on a microarray. *BMC Bioinform.* **2007**, *8*, 220. [[CrossRef](#)]
112. Song, L.; Langfelder, P.; Horvath, S. Comparison of co-expression measures: Mutual information, correlation, and model based indices. *BMC Bioinform.* **2012**, *13*, 328. [[CrossRef](#)] [[PubMed](#)]
113. Yin, L.; Deng, Y. Toward uncertainty of weighted networks: An entropy-based model. *Phys. A Stat. Mech. Its Appl.* **2018**, *508*, 176–186. [[CrossRef](#)]
114. Ferrer-i-Cancho, R.; Dębowski, L.; del Prado Martín, F.M. Constant conditional entropy and related hypotheses. *J. Stat. Mech. Theory Exp.* **2013**, *2013*, L07001. [[CrossRef](#)]
115. Press, W.H. *Numerical Recipes C Diskette V1. 3: Format IBM PC, PS/2, and Compatibles (DOS)*; Cambridge University Press: Cambridge, UK, 1988.
116. Watanabe, S. Information Theoretical Analysis of Multivariate Correlation. *IBM J. Res. Dev.* **1960**, *4*, 66–82. [[CrossRef](#)]
117. Smith, R. A mutual information approach to calculating nonlinearity. *Stat* **2015**, *4*, 291–303. [[CrossRef](#)]
118. Thomas, M.; Joy, A.T. *Elements of Information Theory*; Wiley-Interscience: New York, NY, USA, 2006.
119. Arias-Michel, R.; García-Torres, M.; Schaerer, C.; Divina, F. Feature Selection Using Approximate Multivariate Markov Blankets. In Proceedings of the Automata, Languages and Programming, Rome, Italy, 12–15 July 2016; pp. 114–125.
120. Urbanowicz, R.J.; Meeker, M.; La Cava, W.; Olson, R.S.; Moore, J.H. Relief-based feature selection: Introduction and review. *J. Biomed. Inform.* **2018**, *85*, 189–203. [[CrossRef](#)]
121. Todorov, A. An overview of the RELIEF algorithm and advancements. In *Statistica I Approaches to Gene X Environment Interactions for Complex Phenotypes*; MIT Press: Cambridge, MA, USA, 2016.
122. Kononenko, I.; Robnik-Šikonja, M.; Pompe, U. ReliefF for estimation and discretization of attributes in classification, regression, and ILP problems. *Artif. Intell. Methodol. Syst. Appl.* **1996**, 31–40.
123. Kononenko, I.; Šimec, E.; Robnik-Šikonja, M. Overcoming the Myopia of Inductive Learning Algorithms with RELIEFF. *Appl. Intell.* **1997**, *7*, 39–55. [[CrossRef](#)]
124. Song, Y.; Si, W.; Dai, F.; Yang, G. Weighted ReliefF with threshold constraints of feature selection for imbalanced data classification. *Concurr. Comput. Pr. Exp.* **2020**, *32*, 5691. [[CrossRef](#)]
125. Urbanowicz, R.J.; Olson, R.S.; Schmitt, P.; Meeker, M.; Moore, J.H. Benchmarking relief-based feature selection methods for bioinformatics data mining. *J. Biomed. Inform.* **2018**, *85*, 168–188. [[CrossRef](#)] [[PubMed](#)]
126. Wolter, K.M.; Wolter, K.M. *Introduction to Variance Estimation*; Springer: Berlin/Heidelberg, Germany, 2007; p. 53.
127. Hoerl, A.E.; Kennard, R.W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **1970**, *12*, 55–67. [[CrossRef](#)]
128. Ogotu, J.O.; Schulz-Streeck, T.; Piepho, H.-P. Genomic selection using regularized linear regression models: Ridge regression, lasso, elastic net and their extensions. *BMC Proc.* **2012**, *6*, S10.
129. Moody, G.B.; Mark, R.G.; Goldberger, A.L. PhysioNet: A Web-based resource for the study of physiologic signals. *IEEE Eng. Med. Biol. Mag.* **2001**, *20*, 70–75. [[CrossRef](#)]
130. Wanhammar, L.; Saramäki, T. *Digital Filters Using MATLAB*; Springer Science and Business Media LLC: Cham, Switzerland, 2020.

131. Esmaili, A.; Kachuee, M.; Shabany, M. Nonlinear Cuffless Blood Pressure Estimation of Healthy Subjects Using Pulse Transit Time and Arrival Time. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 3299–3308. [[CrossRef](#)]
132. Cox, R. Regional variation of series elasticity in canine arterial smooth muscles. *Am. J. Physiol. Circ. Physiol.* **1978**, *234*, H542–H551. [[CrossRef](#)] [[PubMed](#)]
133. Nürnberger, J.; Dammer, S.; Saez, A.O.; Philipp, T.; Schäfers, R.F. Diastolic blood pressure is an important determinant of augmentation index and pulse wave velocity in young, healthy males. *J. Hum. Hypertens.* **2003**, *17*, 153–158. [[CrossRef](#)] [[PubMed](#)]
134. Clifford, G.D.; Scott, D.J.; Villarreal, M. User guide and documentation for the MIMIC II database. *MIMIC-II Database Version 2009*, *2*, 95.
135. Kachuee, M.; Kiani, M.M.; Mohammadzade, H.; Shabany, M. Cuffless Blood Pressure Estimation Algorithms for Continuous Health-Care Monitoring. *IEEE Trans. Biomed. Eng.* **2016**, *64*, 859–869. [[CrossRef](#)] [[PubMed](#)]
136. Li, Y.; Chen, X.; Zhang, Y.; Deng, N. Noninvasive continuous blood pressure estimation with peripheral pulse transit time. In Proceedings of the 2016 IEEE Biomedical Circuits and Systems Conference (BioCAS), Shanghai, China, 17–19 October 2016; pp. 66–69.
137. Pinto, E. Blood pressure and ageing. *Postgrad. Med. J.* **2007**, *83*, 109–114. [[CrossRef](#)]
138. Miao, F.; Liu, Z.-D.; Liu, J.-K.; Wen, B.; He, Q.-Y.; Li, Y. Multi-Sensor Fusion Approach for Cuff-Less Blood Pressure Measurement. *IEEE J. Biomed. Health Inform.* **2019**, *24*, 79–91. [[CrossRef](#)]
139. Huynh, T.H.; Jafari, R.; Chung, W.-Y. Noninvasive Cuffless Blood Pressure Estimation Using Pulse Transit Time and Impedance Plethysmography. *IEEE Trans. Biomed. Eng.* **2019**, *66*, 967–976. [[CrossRef](#)]
140. Slapničar, G.; Mlakar, N.; Luštrek, M. Blood Pressure Estimation from Photoplethysmogram Using a Spectro-Temporal Deep Neural Network. *Sensors* **2019**, *19*, 3420. [[CrossRef](#)] [[PubMed](#)]
141. Ramakrishna, P.; Nabeel, O.M.; Joseph, J.; Sivaprakasam, M. Cuffless Blood Pressure Estimation Using Features Extracted from Carotid Dual-Diameter Waveforms. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; Volume 2020, pp. 2719–2722.