Title	A Prediction Model of Blood Pressure for Telemedicine
Author(s)	Kwong, EWY; Wu, H; Pang, GKH
Citation	Health Informatics Journal, 2016
Issued Date	2016
URL	http://hdl.handle.net/10722/229175
Rights	

Title: A Prediction Model of Blood Pressure for Telemedicine

Authors:

Enid Wai-Yung KWONG
School of Nursing, Tung Wah College
Email: enidkwong@twc.edu.hk

Hao WU
Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong.
Email: haowu@eee.hku.hk

Grantham Kwok-Hung PANG
Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong.
Email: gpang@eee.hku.hk

Correspondence should be addressed to:
Hao WU
Phone: +852 - 5182 - 5979
E-mail: haowu@eee.hku.hk
Address: RM510, Chow Yei Ching Building, Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong.

ABSTRACT

Background:

This paper presents a new study based on a machine learning technique, specifically an artificial neural network, for predicting systolic blood pressure through the correlation of variables (age, BMI, exercise level, alcohol consumption level, smoking status, stress level, and salt intake level).

Materials and Methods:

The study was carried out using a database containing a variety of variables/factors. Each database of raw data was split into two parts: one part for training the neural network and the remaining part for testing the performance of the network. Two neural network algorithms, back-propagation and radial basis function, were used to construct and validate the prediction system.

Results:

According to the experiment, the accuracy of our predictions of systolic blood pressure values exceeded 90%. Our experimental results show that artificial neural networks are suitable for modeling and predicting systolic blood pressure.

Conclusions:

This new method of predicting systolic blood pressure helps to give an early warning to adults, who may not get regular blood pressure measurements that their blood pressure might be at an unhealthy level. Also, because an isolated measurement of blood pressure is not always very accurate due to daily fluctuations, our predictor can provide the predicted value as another figure for medical staff to refer to.

Keywords: Systolic blood pressure; artificial neural network; prediction; hypertension; machine learning; telemedicine.

1. Introduction

Hypertension is a major global public health issue. According to the National Institutes of Health (NIH), around one in three adults in the US have hypertension [1]. It is a chronic medical condition characterized by persistent high blood pressure (BP) with systolic and diastolic BP readings of higher than 140mmHg and 90 mmHg, respectively [2].

1.1 Health consequences

Untreated high BP or hypertension may damage major organs of the body, including the heart, kidneys, and brain. Recent studies have shown that hypertension is related to many diseases, such as stroke, loss of consciousness, memory loss, heart disease, cardiovascular disease, and damage to eyes and kidneys [3-8]. Another study has also shown that people who have consistent high blood pressure in their 20s are more likely to develop atherosclerosis by the time they reach middle age [9]. Most people with high blood pressure or hypertension often have no signs or symptoms, even if their blood pressure readings have already reached an emergent level. Thus, they may notice nothing for years, even as their heart, blood vessels, kidneys, brain, or other body parts are damaged by high blood pressure. The early detection or prediction of consistently high blood pressure to minimize the risk of hypertension and its complications is the first step in hypertension care. However, the fact is that many people, especially young and middle-aged adults, do not undergo regular monitoring of their blood pressure in order to prevent hypertension.

1.2. Factors associated with high blood pressure

According to the World Health Organization [1], high blood pressure is very often associated with such health factors as old age, tobacco use, stress, a lack of physical activity, the harmful use of alcohol, and obesity. Below is a discussion of the factors associated with high blood pressure:

Age: Sparrow et al. [10] used multiple linear regression analyses to show that the systolic blood pressure (SBP) of a person aged over 50 increases more rapidly than that of someone between the ages of 20 and 39.

Male/Female: According to a survey [11], the blood pressure of men is usually higher than that of women of the same age. In fact, men are at a higher risk of developing hypertension and cardiovascular disease than premenopausal women of the same age.

BMI: Obesity, as indicated by body mass index (BMI), has become a serious public health issue all over the world. Golino et al. [12] carried out a study to predict blood pressure by the machine learning method and concluded that body mass index, waist circumference, and waist-hip ratio have a relationship of correlation with heart diseases such as hypertension, shock, and cardiovascular diseases.

Alcohol consumption: According to Criqui et al. [13], alcohol consumption seems to be significantly related to systolic and diastolic blood pressure for both men and women. Hartung et al. [14] obtained a similar result, concluding that both males and females blood pressure readings are related to alcohol consumption. They also found that the blood pressure of a regular consumer of medium quantities of alcohol is much lower than that of one who drinks excessive amounts daily, which shows that the relationship between blood pressure and alcohol intake quantity is not a linear one.

Exercise level: Many studies [14] have shown that a person's blood pressure drops in response to heightened physical activity. A lack of physical activity also increases a person's risk of becoming overweight, raising the likelihood that the individual will develop hypertension.

Stress level: The American Heart Association (AHA) has pointed out that although stress is not a confirmed risk factor for high blood pressure, many studies on the subject are currently being carried out. In a stressful situation, people's bodies react by releasing stress hormones into the blood, which raises their blood pressure [15].

Salt intake level: A study by Denton et al. [16] showed that eating too much salt contributes to a significant rise in both systolic and diastolic blood pressure.

Smoke status: Tobacco and exposure to secondhand smoke have a significant effect on blood pressure. Experiments have shown that a person's blood pressure will temporarily increase a few minutes after that person starts smoking [17, 18].

1.3. Model to predict blood pressure

In this paper, a new model for predicting blood pressure is presented. It will allow health service providers to predict a person's blood pressure based on that person's health-related information. This prediction model is not meant for patients who are undergoing treatment, as such patients would already be having their blood pressure measured using standard equipment. The model is designed for people who are not getting regular blood pressure measurements. For example, in a telemedicine consultation, the doctor may want to predict whether a person is at risk of high blood pressure. The information on that person, together with this model, would provide a prediction of that person's blood pressure before an actual measurement is made. In

another situation, a person may want to review his or her health status using an online computer program such as an expert system. The expert system would gather all of the relevant data and health-related information on that person under consultation, and give predictions on the person's potential risk of developing various conditions, including high blood pressure. The model in this paper would be a valuable tool for such predictions, as it provides a preliminary assessment of a person's blood pressure.

Some studies [20, 21] have indicated that there are people who are at risk of developing hypertension, yet do not get their blood pressure checked regularly. Therefore, it is desirable to develop a method that can predict blood pressure using factors relating to a person's age and lifestyle (BMI, exercise level, alcohol consumption level, smoking status, stress level, and salt intake level). The method discussed in this paper is a method for predicting the systolic blood pressure of a person when given factors associated with blood pressure. It is not intended to replace actual measurements, but simply to provide an early warning of an individual's risk of developing hypertension and cardiovascular diseases.

1.4. Literature Review

This paper proposes a new model that will allow health personnel or doctors to know the predicted value of an individual's blood pressure based on that person's health-related information, obtained through the use of a neural network. J. C. D. Convay et al. [7] proposed a blood plasma model to evaluate the relationship between the magnesium ions present in a person's blood plasma and that person's systolic blood pressure using an artificial neural network. Their result showed that the neural network can uncover the complexity of the relationship between the input variables and help the researcher to understand the role of magnesium in SBP and glucose. However, the objective of their study was to determine the

relationship between the magnesium ions in blood plasma and systolic blood pressure, which differs from the aim in our paper.

H. F. Golino et al. [12] carried out a study to predict a person's blood pressure using a classification tree, which is a machine learning technique. The variables used for the prediction were physical conditions, such as body mass index, waist and hip circumference, and waist-to-hip ratio. The machine learning technique that they used differs from the technique used in this paper, which is the artificial neural network. Also, they only used the physical condition of a person to predict that person's blood pressure, whereas we also apply some health-related information of a person (e.g., age, stress, and exercise levels) in the prediction.

A. Kupusinac et al. [19] presented a study using an artificial neural network to estimate a person's cardio-metabolic risk (CMR), with the aim of achieving early prevention of atherosclerosis and cardiovascular diseases. No prediction of blood pressure was carried out. They used gender, age, waist-to-height ratio, body mass index, and blood pressure as the primary risk factors; and HDL-, LDL-, and total cholesterol levels, triglycerides, glycaemia, fibrinogen, and uric acid as the secondary risk factors. The output cmr-coefficient is a relative measure of CMR in the interval [0, 1]. They obtained a result by estimating CMR with a best accuracy of 82.76% after analyzing different parameters of the artificial neural network. In their experiment, they tested the trained artificial neural network using a training dataset and an unknown dataset. However, the goal of their research was to estimate a value that would indicate a person's risk of developing cardio-metabolic disease, which is different from the objective in this paper of predicting the systolic blood pressure.

A review of the impact of telemonitoring blood pressure on hypertension outcomes was given in [20]. Directly measuring BP at home or in the office is crucial for controlling and

managing BP. For patients with diabetes, monitoring blood pressure and blood glucose at home is mandatory for self-managing the disease [21]. As discussed in [21], this may appear to be a simple process in daily life, but in practice there could be all kinds of obstacles and complications involved in uploading the readings from the measurement device. However, this paper focuses on the indirect gathering of information on BP by predictions based on other physical attributes of the patient.

A thorough review indicated that no studies have been carried out based on the same idea as that in our model. Hence, the concept presented in this paper is a new one in the field of telemedicine. As a tool for predicting blood pressure, it can be used as a component of an expert system related to medical consultations.

1.5. Clinical Relevance

It must be noted that a person's normal blood pressure range should correspond to that person's age group [10]. It is well known that blood pressure has a daily pattern that can vary across several BP readings within a day. Blood pressure is usually lower when a person is asleep, rising during the day and reaching a peak in mid-afternoon. Blood pressure values then begin to drop again, and during sleep can be 20% lower than during the day.

The clinical relevance of the method developed in this paper is explained as follows. Suppose the blood pressure of a person is not available or the person is at a remote location. Figure 1 shows an artificial neural network that can be used to predict the systolic blood pressure of a person based on a database used in this study. In this example, the person has to input some values related to age, gender, height, weight, stress, and exercise levels. The neural network will give a prediction of that person's systolic blood pressure. Assuming that the predicted value is larger than the hypertension alarm value, the user should be given a warning.

For example, if the predicted value is 150mmHg, which is more than the systolic BP of 142mmHg for a person aged 52, the user should receive an alert.
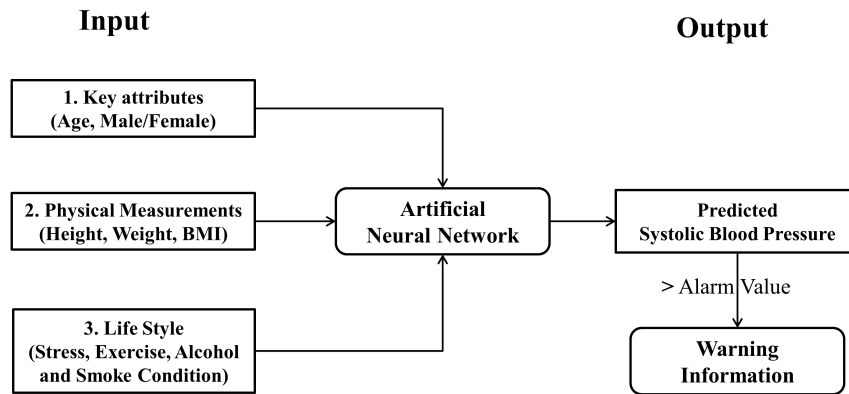

Fig. 1. Predicting without Measuring

Another scenario on the usefulness of the blood pressure predictor is as follows. A person provides all of the relevant information for the neural-network-based blood pressure predictor, and his blood pressure measurement is also taken (Fig. 2). Suppose that the measured value seems normal, but the predicted value is alarming. A recommendation would then be given to the person to double-check his blood pressure. For example, if the measured value was 138mmHg for a person aged 52 but the predictor gives a value of 150mmHg, the relatively low value of his systolic blood pressure measurement could be inaccurate. In such a situation, the person is advised to have further measurements of his blood pressure taken.

Input   **Output**

```
┌─────────────────┐
│ 1. Key attributes│
│ (Age, Male/Female)│
└─────────────────┘

┌─────────────────┐      ┌──────────────┐      ┌──────────────────┐
│ 2. Physical     │      │  Artificial  │      │   Predicted      │
│ Measurements    │─────▶│    Neural    │─────▶│ Systolic Blood   │
│ (Height, Weight,│      │   Network    │      │   Pressure       │
│  BMI)           │      └──────────────┘      └──────────────────┘
└─────────────────┘

┌─────────────────┐
│ 3. Life Style   │
│ (Stress,        │                              ◇ Comparison ◇
│ Exercise, Alcohol│
│ and Smoke       │                                  │ Condition A
│ Condition)      │                                  ▼
└─────────────────┘                          ┌──────────────┐
                                              │   Warning    │
┌─────────────────┐                           │ Information  │
│ 4. Measured     │                           └──────────────┘
│ Systolic Blood  │
│ Pressure        │
└─────────────────┘
```

Condition A:
If Measured SBP seems normal, but Predicted SBP is at an alarming level, the person should double check on his SBP
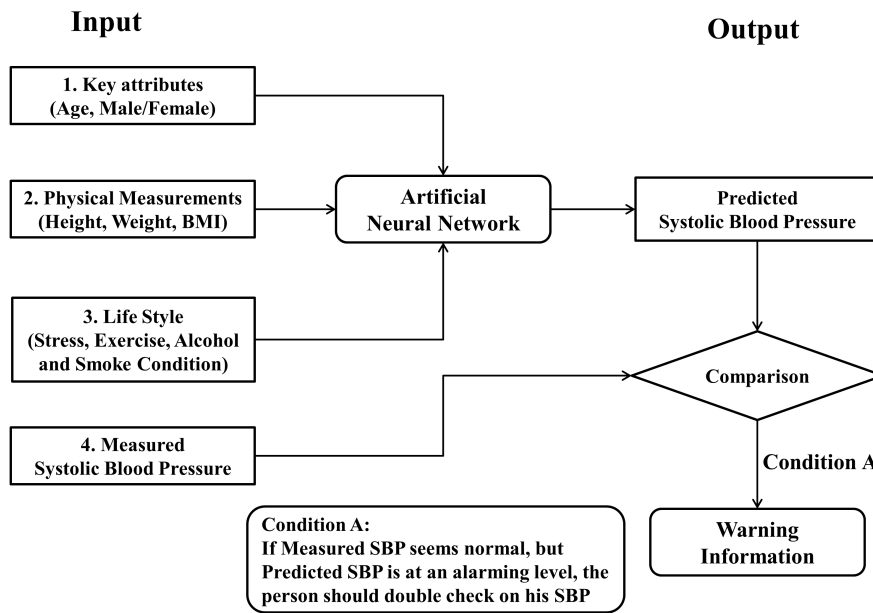
Fig. 2. Predicting with Measurement

In this paper, the machine learning method of an artificial neural network is used to determine the relationship between the correlated variables (age, gender, body mass index, stress level, exercise level, alcohol consumption level, smoking status, and salt intake level) and the systolic blood pressure value. A trained neural network would then be used to predict the systolic blood pressure when the input variables are provided.

2. Materials and methods

2.1. Artificial neural network

The artificial neural network (ANN) is based on an intelligent computational model and uses a computer network system to simulate the biological neural network. It consists of an input layer of source information, at least one middle or hidden layer of computational neurons, and an output layer of computational neurons [22, 23, 24]. The ANN has the ability to discover

the existence of a nonlinear relationship between its inputs and outputs. It has played a major role in many scientific and industrial applications, such as time series prediction, pattern recognition, decision making, load forecasting, and event prediction [19, 25, 26, 27, 28, 29]. In this paper, we aim to use neural networks to determine the correlation between the input variables and the output, which is systolic blood pressure. Essentially, we would like to use an artificial neural network as a tool to predict the value of a person's blood pressure, given some health-related measurements and personal data.

For example, Fig. 3 shows ANN structures using different databases in this study. Based on the database, an ANN structure with seven inputs (age, body mass index, stress level, exercise level, alcohol consumption level, smoking status, and salt intake level) and four hidden nodes is shown in Fig. 3. The common objective of this structure is to predict systolic blood pressure based on the different input variables.
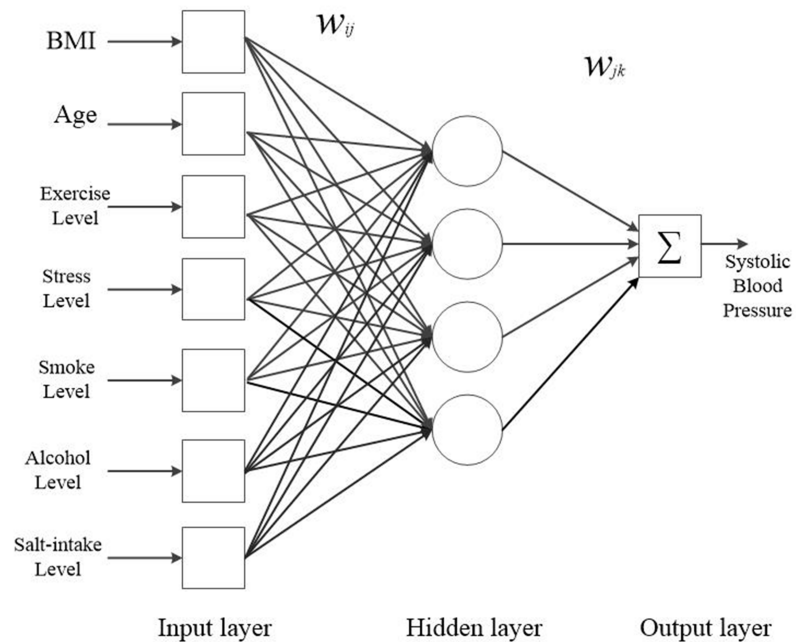

Fig. 3. Structure of an Artificial Neural Network

The strength of a neural network is its learning capability. A problem can be solved by first training the neural network using some training examples or instances of input-output pairs. Two well-known algorithms for the training of neural networks are the back-propagation (BP) algorithm and the radial basis function (RBF) network. The interconnection weights between the different layers in an ANN will be obtained at the end of the training session.

Back-propagation Neural Network [22, 23, 30, 31, 32]

The back-propagation (BP) neural network is a typical architecture of a multi-layered feed-forward neural network. A typical BP neural network usually consists of layers of neurons, and the objective is to train the network weights so as to minimize the mean-square error of the network output. Fig. 3 shows the architecture of a typical BP neural network with seven inputs, one hidden layer with four hidden nodes, and one output value.

Radial Basis Function Neural Network [33, 34]

Our radial basis function (RBF) neural network is composed of one hidden layer. In an RBF network, the hidden layer can implement better nonlinear mapping from the data space to the feature space. The most common basis function chosen is a Gaussian function, which means that the closer the input is to the center of the Gaussian function, the larger will be the response of the node. In stage one of the training of the RBF network, the self-organizing feature map (SOFM) was used as a clustering algorithm to determine the weights between the input and the hidden layer. At each hidden node, the basis function used was a Gaussian function. In the second stage, the weights between the hidden and output layer were determined by a least-mean-square algorithm.

2.2. Statistics

The data were collected from the health and body conditions of 498 people, and included systolic blood pressure (SBP), gender, age, BMI, smoking status, exercise level, alcohol consumption level, stress level, and salt intake level [35]. These variables are believed to be correlated with a person's blood pressure reading. The descriptions of the variables are shown in Table 1. Among the 498 cases, 236 were male, accounting for 47% of the data, while the remaining 262 cases were female.

Due to the obvious influence of gender difference, the author split the data into male and female sets, respectively.

Table 1. Description of the variables

Variable	Description
Systolic Blood Pressure (SBP)	Continuous Variable (in mm Hg)
Gender	Binary Variable (M – Male; F – Female)
Age	Continuous Variable (in year)
Body Mass Index (BMI)	Continuous Variable (in kg/m2)
Smoking Status	Binary Variable (Yes, No)
Exercise Level	Categorical Variable (Low, Medium, High)
Stress Level	Categorical Variable (Low, Medium, High)
Alcohol Level	Categorical Variable (Low, Medium, High)
Salt Intake Level	Categorical Variable (Low, Medium, High)

3. Results

The experiments were carried out by two different algorithms of artificial neural networks: back-propagation (BP) and radial basis function (RBF). The BP algorithm was first used for the training of an ANN. The output was the predicted value of the systolic blood pressure. The inputs were the factors used for the prediction (age, gender, body mass index, stress level, exercise level, alcohol consumption level, smoking status, salt intake level), and the number of hidden nodes varied from 1 to 10.

The RBF algorithm was also used for the training of an ANN, which mapped the relationships between systolic blood pressure (SBP) and the input variables belonging to different clusters. The number of naturally formed clusters was determined, and these then provided the weights between the input and hidden layers. The maximum number of clusters varied from 1 to 10. In the second stage, the weights were determined by an iterative least-mean-square (LMS) algorithm.

The flowchart of the algorithm is shown in Fig. 4. The experiment was carried out and repeated 50 times for each problem instance, and the average value of the results was presented. To evaluate the ANN with different hidden nodes/clusters, the algorithm varied the number of nodes/clusters from 1 to 10.
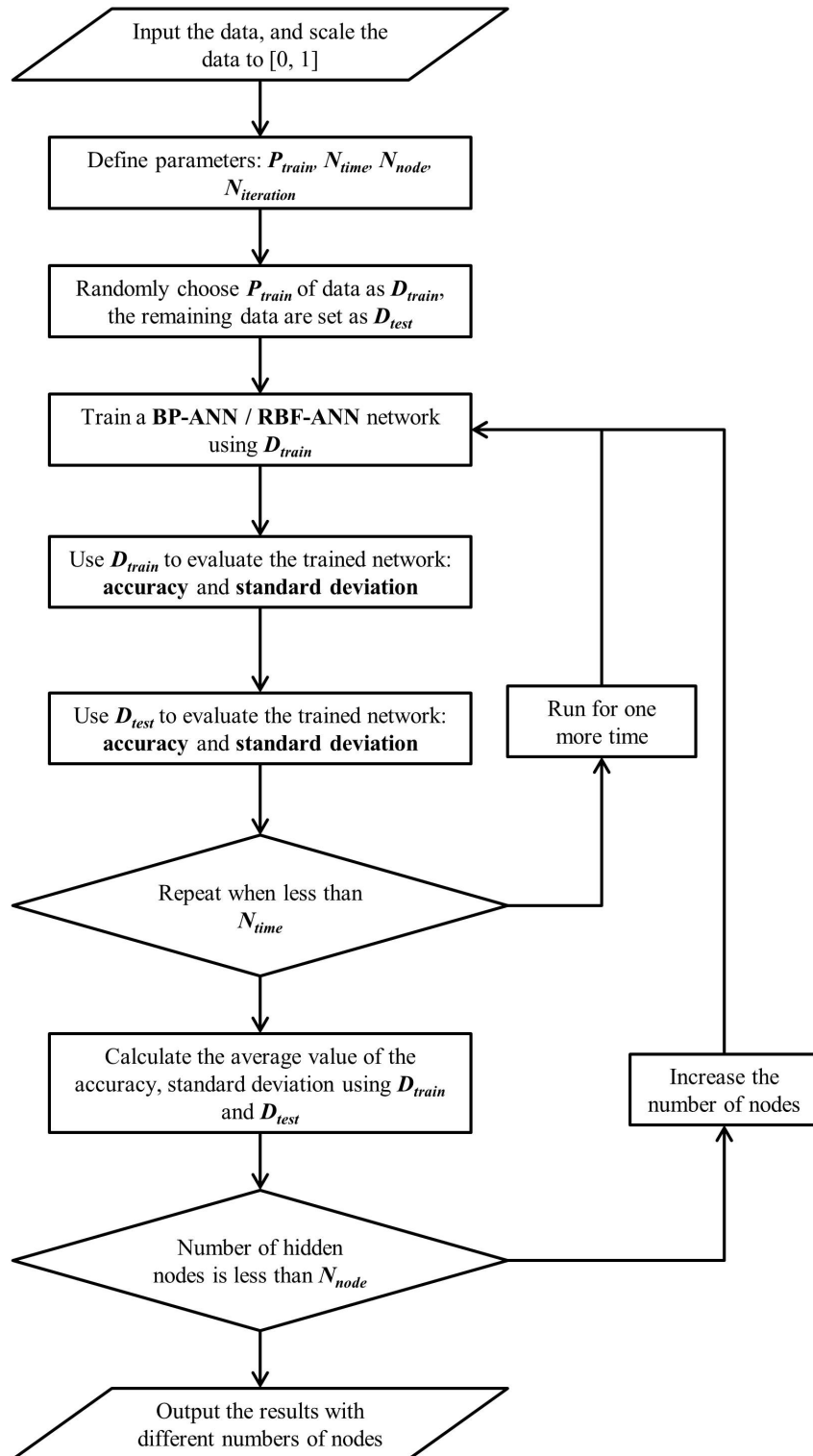
Fig. 4. Flowchart of the ANN algorithm

It is very important to emphasize that the testing data was kept separate from the training data. A certain percentage (60%, 70%, 80%, or 90%) of the data from the database was used for training, and the rest was used for testing the performance of the trained neural network. The predicted systolic blood pressure value from the neural network was compared with the value in the database. The average accuracy, defined below, was recorded, and the standard deviation was also obtained.

$$\text{Accuracy} = 1 - \frac{\left|\text{Predicted Value} - \text{Measured Value}\right|}{\text{Measured Value}} \times 100\%$$

The blood pressure prediction model was coded in Matlab R2014b and run on an Intel Core 2 Quad 2.83GHz PC with 8GB of RAM. The computation times of the experiments are presented in the following parts.

3.1. Back-propagation ANN

In our experiment, we have already investigated the parameters relating to the performance of the neural network, such as different hidden neurons of ANN, and different proportions for training and testing. The experiment was repeated 50 times, and the average of the results on accuracy and standard deviation was obtained. The trained neural network was tested by two sets of data: the first set was the data used to train the ANN, and the second set was the data reserved for testing. The experiment was carried out using the MATLAB neural network toolbox. The algorithm of the toolbox was set to trainlm and the learning rate was 0.3. One hidden layer of neurons was used initially in the BP algorithm, and an architecture of two hidden layers was also examined later. We examined the issue of the maximum number of iterations in the training of the BP neural network. We found that the training settled down within 200 iterations. Two typical performance graphs are shown in Fig. 5.
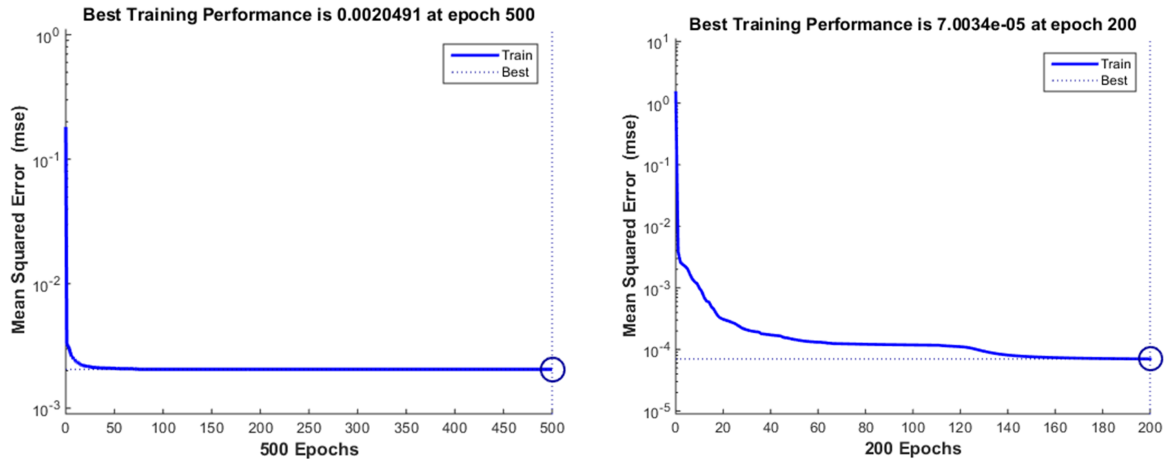
Fig. 5. Two performance graphs of convergence obtained by BP ANN

In our experiment, we allowed the maximum number of iterations to reach 200.

<u>Results based on Males</u>

The ANN used one hidden layer at the beginning of this experiment. Different percentages of the database were used for the training of the BP ANN. As shown in Fig. 6, lines of different color stand for different proportions (60-40, 70-30, 80-20, and 90-10) of training and testing from the dataset. Based on the testing data, it is obvious that the best result was obtained when 80% was used for training and the remaining 20% for testing the performance of the ANN. As the number of hidden nodes increased (from 1 to 10), the performance achieved using the training data kept improving, while the results based on the testing data got worse. This is because when the number of hidden nodes increased, the neural network had a greater degree of freedom to learn the relationship between the inputs and the output based on the training data. Hence, the accuracies based on the training data improved as the number of hidden nodes increased. On the other hand, the use of more hidden nodes decreased the generalization capability of the neural network. When the unseen testing data was presented to the neural network, the prediction accuracies appeared to decrease as the number of hidden nodes

increased. Taking the two graphs on accuracy into consideration, we found that using four hidden nodes gives the best overall result when their averages are calculated.
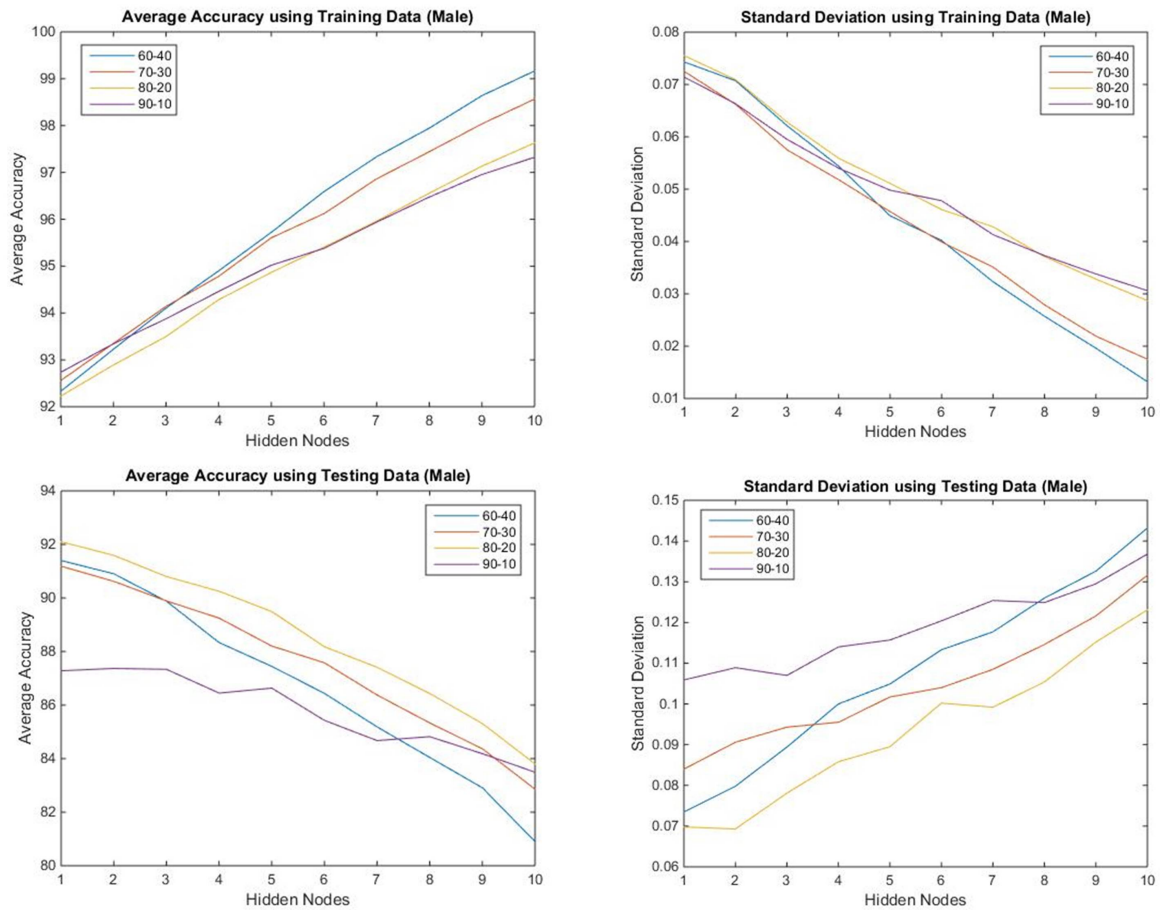

Fig. 6. Performance of trained BP ANN using the male data (one hidden layer)

Results based on Females

A similar experiment was carried out for the female cases in the dataset. Referring to the different proportions on the use of data (Fig.7), and on the accuracy of predictions using the testing data, the use of 80% data for training gave the "best" result. Again, it seemed that more hidden nodes would give better results when using the training data, which means that the ANN had already learned the corresponding relationship of the training data well. However, the trained ANN with more hidden nodes did not perform well when using the testing data.

Taking the average of the two results into consideration, the use of four hidden nodes gave the best result.
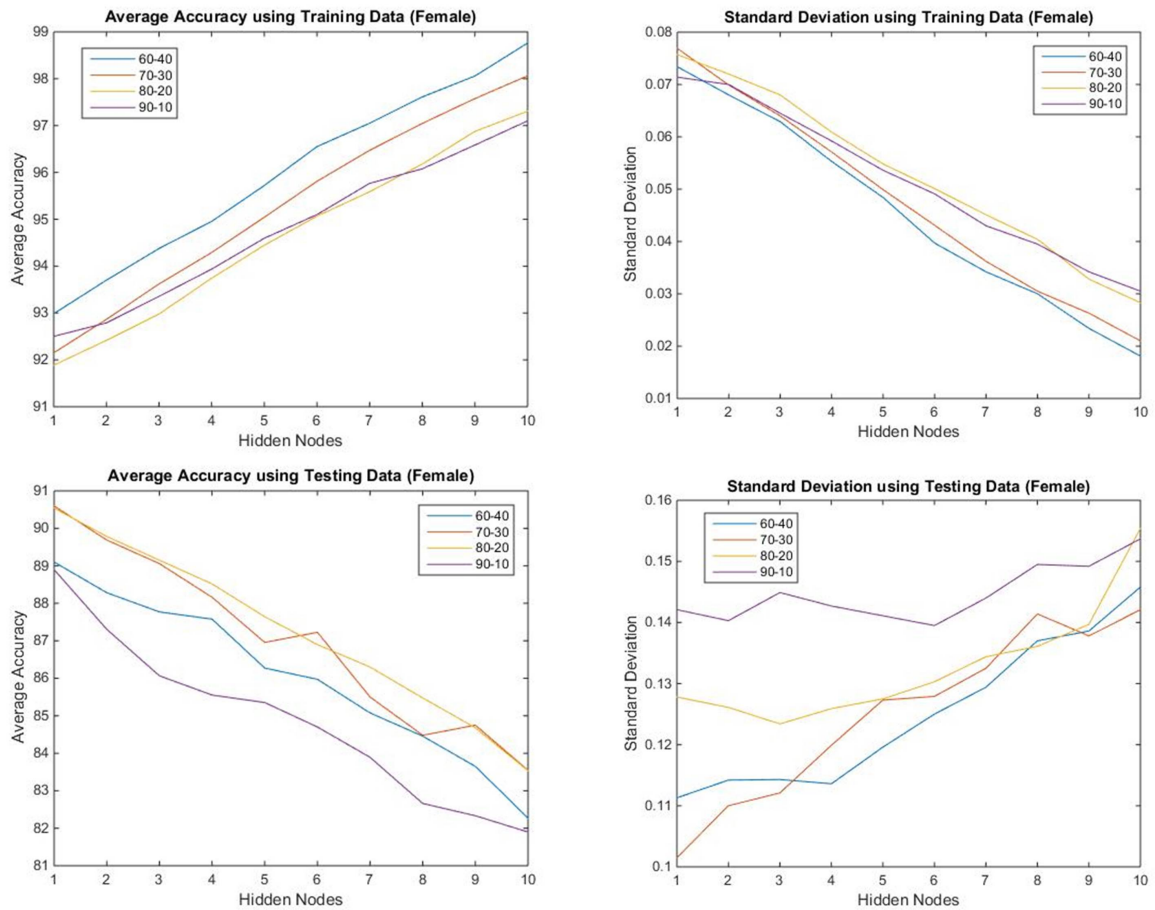

Fig. 7. Performance of trained BP ANN using the female dataset

Results based on both genders

A case combining the male and female datasets was carried out in this section. The experiment was run and repeated 50 times, after which the average accuracy was calculated. The number of hidden nodes varied from 2 to 10. In the case combining both genders, the gender information of each sample was defined as another input in the ANN. In Table 2, the performance of this case was compared with those of cases of datasets separated by gender.

The last row shows that the prediction accuracy with a combined gender dataset is lower than that with male cases.

Table 2. Comparison of the performances of the combined and separated datasets

Number of hidden nodes	2	3	4	5	6	7	8	9	10
Male	91.66%	**91.69%**	91.59%	91.62%	91.40%	91.22%	91.13%	90.94%	90.47%
Combined	90.86%	90.97%	91.02%	91.08%	91.12%	91.15%	91.13%	**91.19%**	91.14%

<u>Results based on two hidden layers</u>

In this part, a BP-ANN with two hidden layers was used to obtain the prediction value of the blood pressure. To test the performance of the network, the results were compared with the experiment using male data with one hidden layer. In this test, 80% of the data were used for training, with the remainder being used for testing. Each experiment was also repeated 50 times, and then the average result was calculated. In the network with two hidden layers, the number of hidden nodes for each layer varied from 2 to 10. As shown in Table 3, the average accuracy using the network with one hidden layer was always higher than the network with two hidden layers. Hence, having two hidden layers did not help the BP-ANN to obtain a predicted blood pressure value in this problem.

Table 3. Comparison of the performances of one and two hidden layers

Number of nodes on each layer	2	3	4	5	6	7	8	9	10
One hidden layer	91.66%	91.69%	91.59%	91.62%	91.40%	91.22%	91.13%	90.94%	90.47%
Two hidden layers	**89.84%**	**89.05%**	**88.04%**	**85.94%**	**84.48%**	**82.57%**	**81.36%**	**81.20%**	**80.91%**

3.2. Radial-Basis Function (RBF) Network

In this experiment, some related parameters were investigated to test the performance of the radial-basis function neural network. As shown in Fig. 3, the RBF network in this paper is a two-layer network whose output forms a linear combination of the basis function computed by the hidden units. The basis functions in the hidden layer produce a localized response to the input, and each hidden unit has a localized receptive field.

Let X and Y be the set of the input and output data, respectively. Let $(x_i, y_i), i = 1, 2, 3, ...,$ $NumT$ be such a training sample, where $NumT$ is the number of training data. Let C be the set of centers/clusters. The algorithm of the RBF network used in this paper is summarized into three steps.

Step 1 – weight initialization: the weights in the hidden layer are determined by a clustering algorithm, and the weights in the output layer are initialized to small random values.

Step 2 – calculation of activation: the activation level h_j of the hidden unit j is calculated by

$$h_j = \exp[-\frac{(X - C_j)^T \cdot (X - C_j)}{2\sigma^2}].$$

The activation level n_j of an output unit is determined by

$$n_j = \Sigma W_{ij} h_i.$$

Step 3 – weights calculations: in the hidden layer, the self-organizing feature map (SOFM) clustering algorithm is used to update the weights. In the output layer, the weight is learned by a least mean square (LMS) algorithm. The weights are adjusted by

$$W_{ij}(t + 1) = W_{ij}(t) + \Delta W_{ij},$$

where $W_{ij}(t)$ is the weight from the unit i to unit j at the t-th iteration and ΔW_{ij} is the weight adjustment.

The weight change is computed by

$$\Delta W_{ij} = \eta \delta_j h_i \text{,}$$

where η is a trial-independent learning rate and δ_j is the error at the unit j.

$$\delta_j = d_j - n_j \text{,}$$

where d_j is the desired output activation, and n_j is the actual output at the output unit j.

Similar to the BP ANN, different proportions were used for training and different numbers of clusters were carried out in the experiment. However, the radial basis function helps the kernel function of ANN rather than the optimization algorithm. The experiment was repeated 50 times, and the average of the results in terms of accuracy and standard deviation was calculated. Again, the trained network was tested based on two sets of data: data used to train the ANN and data reserved for testing. The learning rate of this algorithm was set to 0.3 and the maximum iteration was set to 100. The network had already settled down with 100 iterations, and there was no advantage to having more iterations.

<u>Results based on Males</u>

In this case of RBF NN, Fig. 8 shows that having more clusters did nothing to improve performance. Our experiment found that five clusters were enough for the problem. Different percentages of the database were used for the training of the RBF ANN. In Fig. 8, lines of different colors stand for different proportions (60-40, 70-30, 80-20, and 90-10) of training and testing data from the dataset. Based on the testing data, it seems that the best result was obtained when 70% was used for training and the remaining 30% for testing the performance of the ANN.
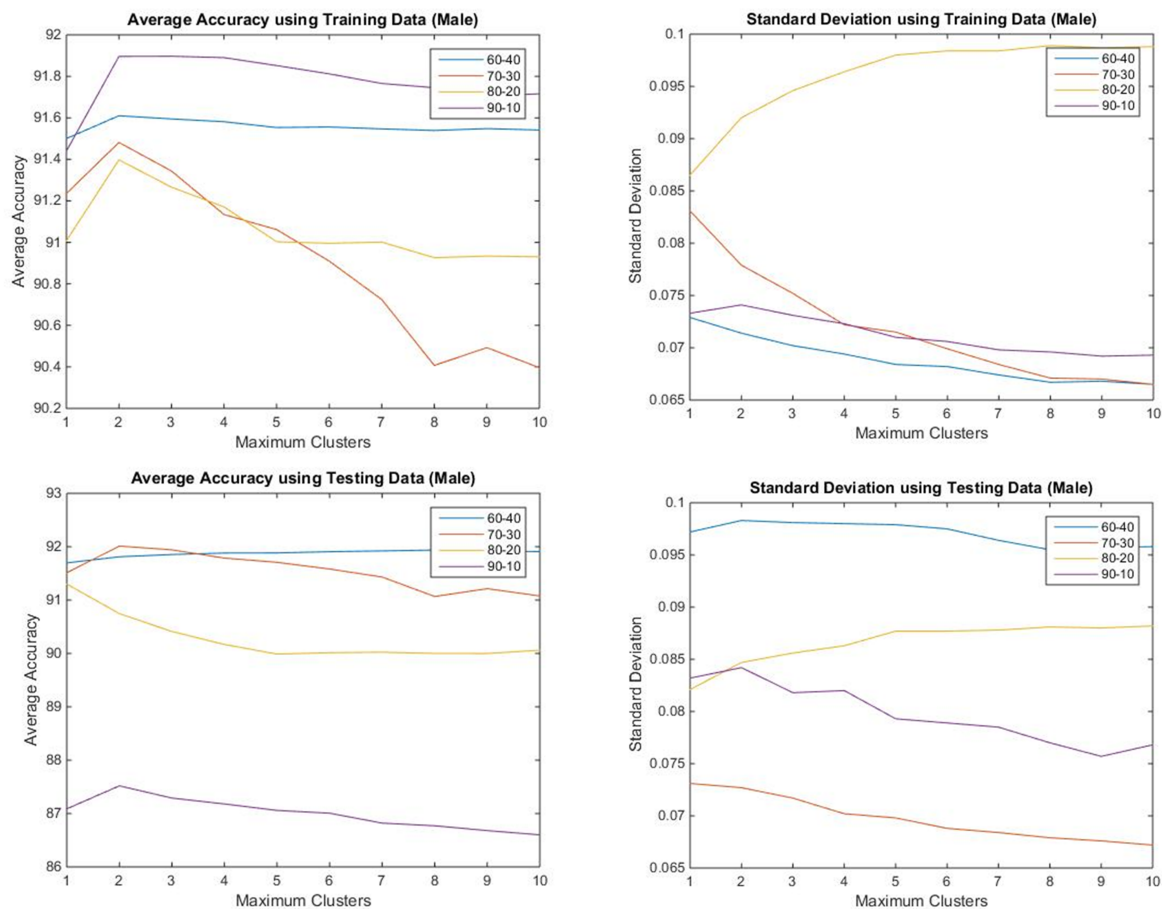
Fig. 8. Performance of trained RBF ANN using the male data

Results based on Females

A similar experiment was carried out for the female cases as shown in Fig.9. Again different percentages (60%, 70%, 80% and 90%) of the database were used for the training of the RBF ANN. Based on the testing data, it seems that the best result is obtained when 90% is used for training and the remaining 10% for testing the performance of the ANN.
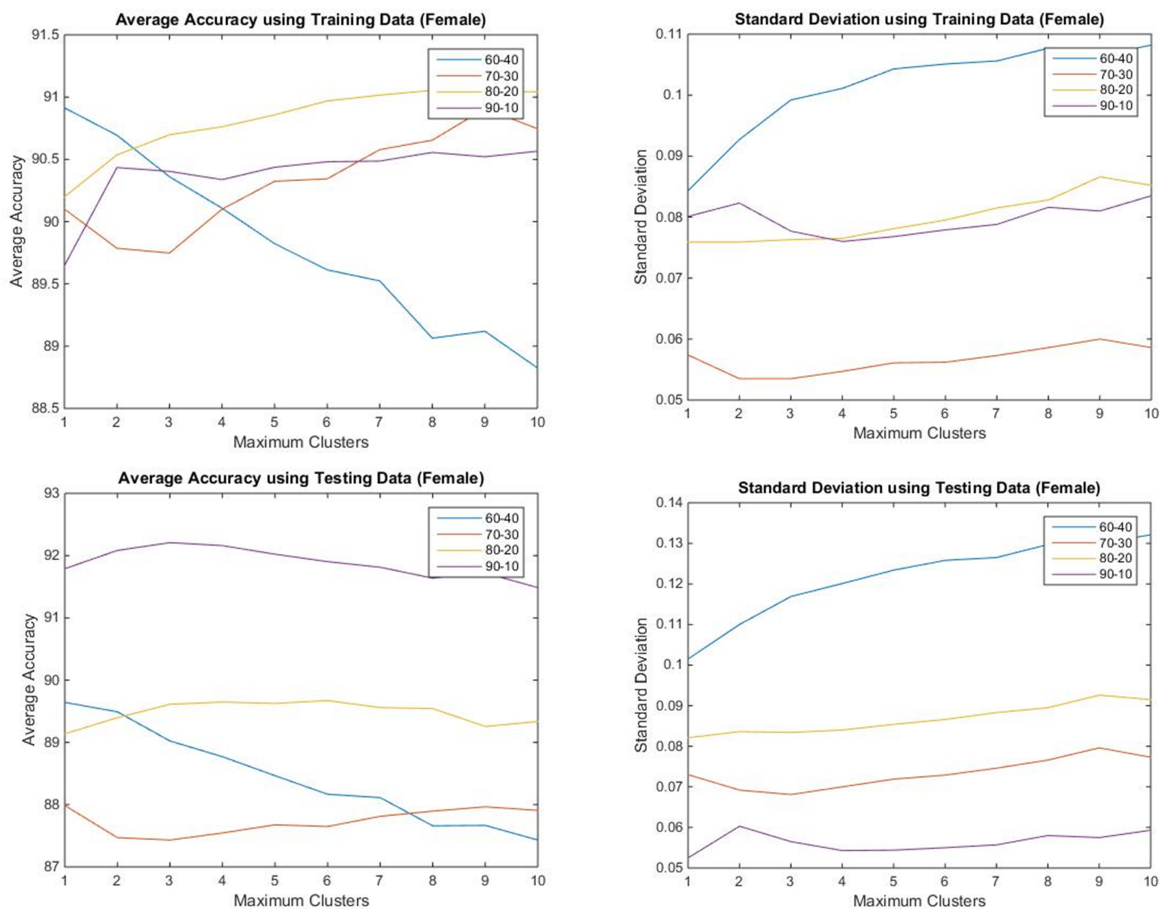
Fig. 9. Performance of trained RBF ANN using the female data

3.3 Summary of the results

The computation times of the experiments are summarized in Table 4. In the experiment, the neural network was tested by the number of hidden nodes/clusters from 1 to 10. Also, each case of experiment was run 50 times to calculate the average accuracy. The average computation time of the ANN was calculated with different proportions of training and testing data. The results showed that the RBF-ANN runs faster than the BP-ANN in obtaining the predicted blood pressure value.

Table 4. Average computation time

Algorithm of ANN	Dataset	Proportion (90/10)	Proportion (80/20)	Proportion (70/30)	Proportion (60/40)
BP - ANN	Male	2.14s	2.10s	2.02s	1.94s
	Female	2.10s	2.04s	1.94s	1.94s
RBF - ANN	Male	0.99s	0.86s	0.68s	0.57s
	Female	1.12s	0.94s	0.74s	0.69s

A summary of the major results is shown in Table 5. It can be seen that the accuracy was around 90% for both male and female cases. The number of nodes used in the middle layer of the neural network was also comparable.

Table 5. Experimental results

Algorithm of ANN	Dataset	Number of hidden nodes	Proportion (training / testing)	Accuracy using trained data	Standard deviation using trained data	Accuracy using testing data	Standard deviation using testing data
BP - ANN	Male	4	80 / 20	90.25%	0.0858	94.28%	0.0559
	Female	4	80 / 20	88.52%	0.1259	93.74%	0.0609
RBF - ANN	Male	5	70 / 30	91.71%	0.0698	91.06%	0.0715
	Female	5	90 / 10	92.02%	0.0544	90.44%	0.0768

4. Discussion

Gender is the only constant factor among the input variables. Other variables, such as BMI and smoking status, change according to a person's lifestyle. Also, using a separated ANN for male cases was shown to lead to better results than using a combined dataset. The database showed that the subjects ranged in age from 18 to 64, with 236 male and 262 female cases. Many of the results have been presented in the previous section. The reliability of the proposed approach was evaluated by comparing the ANN-predicted systolic blood pressure value against the data used for both training and testing. The main criterion for evaluation should be the result obtained based on the testing data, and accuracy of over 90% was obtained.

Both of the artificial neural networks (ANN) studied in this paper had a structure with only one hidden layer. In fact, the use of two hidden layers did not produce better results than the use of one hidden layer. It is also understood in the field of ANN that one hidden layer is sufficient to represent any nonlinear function, provided that enough hidden nodes are used.

We also found that using four hidden nodes in the BP ANN was sufficient for good performance. We attempted to use more hidden nodes in the training of both the BP and RBF networks, but the performance of the prediction was no better than when just a few hidden nodes were used. This observation is not surprising, as it is well known that using more hidden nodes merely introduces more parameters into the predictor (ANN). The neural network tries to "memorize" the training cases rather than to generalize and strengthen its learning capability. This issue of over-training should be avoided; the use of just a few hidden nodes has verified that our approach is correct.

The results suggest that the BP and RBF networks in the current study are comparable. The BP structure provides a default ANN configuration, and the learning effort goes into the training to find the interconnection weights between the layers. However, the RBF is a more refined structure, and the neural network is trying to turn the training data into clusters. Generally speaking, the time required for training is faster for an RBF network than for a BP network.

Limitations of the study

This study has several limitations. First, the size of the database used in the experiments was not very large. In order to obtain better predictions, the database should contain more samples for the training of ANN. Second, some attributes provided in the database, such as stress, exercise, alcohol, and smoking status, were subjective values without detailed and clearer

information. With the use of telemedicine technology [36], many attributes and health-related data, including the patient's BP, can be gathered online. These would form a valuable database for predicting BP. To conclude, further studies using a larger database should be carried out, along with the use of more input variables for the neural network.

5. Conclusions

This paper has proposed a machine learning method to predict systolic blood pressure using the back-propagation (BP) neural network and the radial basis function (RBF) network. The results from this paper indicate that a blood pressure predictor can be developed based on the artificial neural network models that were proposed. The average accuracy of the prediction is over 90%. The results obtained from the BP neural network and RBF network are in agreement.

Our results indicate that the machine learning technique can be an efficient tool for analyzing the relationship between a person's systolic blood pressure and his age and lifestyle factors (BMI, exercise level, alcohol consumption level, smoking status, stress level, and salt intake level). It must be emphasized that this method does not aim to replace the direct measurement of a person's blood pressure. However, the method provides an estimated value in a telemedicine or expert system consultation. It also provides a reference value to compare with the measured value if a measurement is taken. This can contribute to the development of an SBP predictor, which can provide early warnings of the risks of hypertension and cardiovascular disease.

References

[1] World Health Organization, Q&As on hypertension, http://www.who.int/features/qa/82/en/

[2] Difference Between Hypertension and High Blood Pressure, http://www.differencebetween.net/science/health/difference-between-hypertension-and-high-blood-pressure.

[3] Ehret GB, Munroe PB, Rice KM, Bochud M, Johnson AD. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. *J Nature* **2011**; 478: 103-109.

[4] Marik PE, Varon J. Hypertensive crises: challenges and management. *Chest Journal* **2007**; 131: 1949-1962.

[5] Jensen MH, Cichosz SL, Hejlesen OL, Toft E, Nielsen C, Grann O, Dinesen BI. Clinical Impact of Home Telemonitoring on Patients with Chronic Obstructive Pulmonary Disease. *Telemed E Health* **2012**;18:674-678.

[6] American Heart Association, Hypertensive Crisis http://www.heart.org/HEARTORG/Conditions/HighBloodPressure/AboutHighBloodPressure/Hypertensive-Crisis_UCM_301782_Article.

[7] Conway JCD, Lavorato SN, Cunha VF, Belchior JC. Modelling of the relationship between systolic blood pressure and glucose with the magnesium ion present in the blood plasma: an approach using artificial neural networks. *J Health* **2009**; 1: 211-219.

[8] Gursoy G, Ulu S, Acar Y. Relation of resistin with obesity and some cardiovascular risk factors in hypertensive women. *J Res Med Sci* **2012**; 17: 443-447.

[9] Allen NB, Siddique J, Wilkins JT, Shay C, Lewis CE. Blood Pressure Trajectories in Early Adulthood and Subclinical Atherosclerosis in Middle Age. *JAMA* **2014**; 311: 1568.

[10] Sparrow D, Garvey AJ, Rosner B, Thomas JrHE. Factors in predicting blood pressure change. *J Circulation* **1982**; 65: 789-794.

[11] Reckelhoff JF. Gender Differences in the Regulation of Blood Pressure. *Hypertension* **2001**; 37: 1199-1208.

[12] Golino HF, Amaral deLSB. Predicting Increased Blood Pressure Using Machine Learning. *J Obesity* **2014**; 637-635.

[13] Criqui MH, Wallace RB, Mishkel M, Connor EB, Heiss G. Alcohol consumption and blood pressure - the lipid research clinics prevalence study. *Hypertension* **1981**; 3: 557-565.

[14] Hartung GH, Kohl HW, Blair SN, Lawrence SJ, Harrist RB. Exercise tolerance and alcohol intake. Blood pressure relation. *Hypertension* **1990**; 16: 501-507.

[15] American Heart Association, Stress and Blood Pressure, available at; http://www.heart.org/HEARTORG/Conditions/HighBloodPressure/PreventionTreatmentofHig hBloodPressure/Stress-and-Blood-Pressure_UCM_301883_Article.jsp.

[16] Denton D, Weisinger R, Mundy NI, Wickings EJ, Dixson A, Moisson P. The effect of increased salt intake on blood pressure of chimpanzees. *J Nat Med* **1995**; 1: 1009-1016.

[17] American Heart Association, Prevention and Treatment: Tobacco and Blood Pressure, Apr 10, 2014, available at: http://www.heart.org/HEARTORG/Conditions/HighBloodPressure/PreventionTreatmentofHig hBloodPressure/Tobacco-and-Blood-Pressure_UCM_301886_Article.jsp.

[18] Omvik P. How smoking affects blood pressure. *Blood Pressure* **1996**; 5: 71-77.

[19] Kupusinac A, Doroslovački R, Malbaški D, Srdić B and Stokić E. A primary estimation of the cardiometabolic risk by using artificial neural networks. *Computers in Biology and Medicine* **2013**; 751-757.

[20] AbuDagga A, Resnick HE, Alwan M. Impact of Blood Pressure Telemonitoring on Hypertension Outcomes: A Literature Review. *Telemedicine and e-Health* **2010**; 16: 830-838.

[21] Koopman RJ, Wakefield BJ, Johanning JL, Keplinger LE, Kruse RL, Bomar M, Bernt B, Wakefield DS, Mehr DR. Implementing Home Blood Glucose and Blood Pressure Telemonitoring in Primary Care Practices for Patients with Diabetes: Lessons Learned. *Telemedicine and e-Health* **2014**; 20: 253-260.

[22] Haykin S. Neural Network: A Comprehensive Foundation. New Jersey: Pearson Education; 1999.

[23] Kung SY, Mak MW, SH Lin. Biometric Authentication: A Machine Learning Approach. New Jersey: Prentice Hall; 2004.

[24] Graupe D. Principles of Artificial Neural Network. World Scientific; 2013.

[25] Hao DX, Chen R. AFSs-RBF Neural Network for Predicting Earthquake-induced Liquefaction of Light Loam. *Natural Computation (ICNC)* **2010**; 3: 1518-1522.

[26] Kaminski M, Orlowska-Kowalska T. Clustering in optimization of RBF-based neural estimators for the drive system with elastic joint. *Industrial Electronics* **2011**; 1907-1912.

[27] Torbati N, Ayatollahi A, Kermani A. An efficient neural network based method for medical image segmentation. *Computers in Biology and Medicine* **2014**; 76-87.

[28] G.W. Dombi, J.P. Rosbolt and R.K. Severson. Neural network analysis of employment history as a risk factor for prostate cancer. *Computers in Biology and Medicine* **2010**; 751-757.

[29] Tang ACY, Chung JWY, Wong TKS. Digitalizing Traditional Chinese Medicine Pulse Diagnosis with Artificial Neural Network. *Telemedicine and e-Health* **2012**; 18: 446-453.

[30] Niu DX, Shi H, Li JQ, Wei YN. Research on Short-term Power Load Time Series Forecasting model Based on BP Neural Network. *Advanced Computer Control* **2010**; 509-512.

[31] Wilamowski BM. Suitability of fuzzy systems and neural networks for industrial applications. *Optimization of Electrical and Electronic Equipment* **2012**; 1-7.

[32] Huang JC. Exploring the Acceptance of Telecare Among Senior Citizens: An Application of Back-Propagation Network. *Telemedicine and e-Health* **2011**; 17: 111-117.

[33] Yang ZR. A Novel Radial Basis Function Neural Network for Discriminant Analysis. *Neural networks* **2006**; 17: 604-612.

[34] Heimes F, Heuveln BV. The Normalized Radial Basis Function Neural Network. *Systems, Man, and Cybernetics* **1998**; 2: 1609-1614.

[35] Blood Pressure Case Study, http://www.math.yorku.ca/Who/Faculty/Ng/ssc2003/BPMain.htm.

[36] Neumann CL, Menne J, Schettler V, Hagenah GC, Brockes C, Haller H, Schulz EG. Long-Term Effects of 3-Month Telemetric Blood Pressure Intervention in Patients with Inadequately Treated Arterial Hypertension. *Telemedicine and e-Health* **2015**; 21: 145-150.