VARIABLE LEARNING RATE BASED MODIFICATION IN BACKPROPAGATION ALGORITHM (MBPA) OF ARTIFICIAL NEURAL NETWORK FOR DATA CLASSIFICATION

Qamar Abbas[1], Qamar Abbas[2], Farooq Ahmad[3] and Muhammad Imran[4]
[1,2]Computer Department, Iqra University, Islamabad,44000, Pakistan
[3]MathematicsDepartment, Govt. Islamia College Civil Lines, Lahore, Pakistan
[3]Presently: MathematicsDepartment, College of Science, Majmaah Universty, Azzulfi, KSA
[4]Computer Department, Szabist University, Islamabad,44000, Pakistan
itinformations@gmail.com, qamar.bhk@gmail.com, farooqgujar@gmail.com, malikimran110@gmail.com
Corresponding author: Dr. Qamar Abbas: itinformations@gmail.com,

ABSTRACT:: *Learning rate is an important parameter of Backpropagation algorithm (BPA) used to train feed forward artificial neural network. Learning rate has great impact on the training of Backpropagation neural network algorithm. This research work introduces some novel variable schemes of learning rate in BP algorithm. The performances of these new variations are tested over four commonly used and standard benchmark datasets taken from the UCI standard machine learning repository. The algorithm focuses on the classification of non-linear datasets. Most of the researchers have used constant value of learning rate. Change in learning rate changes the training behaviour of BPNN algorithm. Various learning rate schemes of different characteristic introduced in this research work will prove to be valuable addition in BP algorithm. Proposed learning rate schemes are helpful in improving the convergence speed and testing accuracy of BPNN algorithm.*

Keywords: Backpropagation, Learning Rate, Linear increasing, Linear Decreasing, Choatic, Oscillating, Random

1. INTRODUCTION

Artificial neural networks (ANN) consist of parallel processing units having capability of storing experimental knowledge and making it available for use. ANN is a new form of computing, inspired by biological models. ANN is a branch of artificial intelligence where its architecture is modeled in the form of software having similar workings like human brain [1]. ANN is made up of interconnecting artificial neurons. The interconnected neurons are normally called nodes. ANN can be used to solve problems require intelligence or to be used to gain an understanding of biological nervous system without necessarily creating a model of a real biological system [2]. The working of artificial network follows the working to human neural system [3]. Figure 1 show the structure of the human nervous system which consists of neurons.

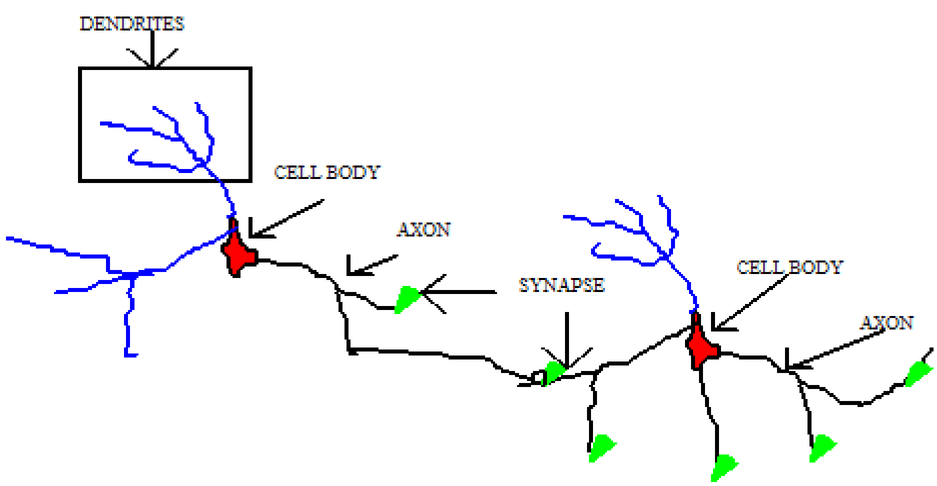


Figure.1 Biological Model of neuron [3]

The dendrites act as the input units of external signals to the neuron and the axon acts as the output unit. The soma (cell body) sums the incoming signals. Signals sent by other neurons are received by the dendrites. The signals are transmitted across synaptic gap by means of a chemical process. Incoming signals are modified by the action of chemical transmitter. This process is similar to weight updating process in ANN. When sufficient input is received then the signal is transmitted to other cells over its axon this process is known as fires. It is often supposed that either cell will fire or not. Axons are the data transmission paths. Similar summation firing is applied in neural networks [3].

Backpropagation algorithm is one of the well-known algorithms of artificial neural network. The pattern learning of ANN is carried out by the weight using in architecture of ANN. The main issue with backpropagation algorithm is the slow convergence speed and local optima problem[4,5,6,7,8] The updation of these weights is important that includes the old value of weigh, the current pattern the network error and learning rate. The old value of weight, the current pattern and the network error have fix values while the value of learning rate is user specified. Most of the researcher uses constant value of leaning rate that may not be less value able for a variety of pattern recognition application. The variable leaning rate will be helpful in escaping from local optima problem and improving the convergence speed of BP algorithm.

Learning rate controls the step size of BPNN algorithm [1]. The small step size results in slow convergence or it can stuck the learning process in local optima problem and big value of learning rate may skip the optimum value that will result in poor learning or overtraining of BPNN algorithm [8]. The main objective of this research work is to introduce such variable learning schemes that may use the combination or some order that uses the small value, big value and values in between small value and big value of learning rate. The combination of various learning rate values will prove to be helpful in increasing the convergence speed of BPNN algorithm and learning accuracy of BPNN algorithm.

This research work introduces five variable learning rate schemes in BPNN algorithm. These variable learning rates will take various values of learning rate from a specific interval of learning rate. The learning rate schemes include learning rate controls the step size of BPNN algorithm. The small step size results in slow convergence or it can stuck the learning process in local optima problem and big value of learning rate may skip the optimum value that will result in poor learning or overtraining of BPNN algorithm [9]. The main objective of this research work is to introduce such variable learning schemes that may use the combination or some order that uses the small value, big value and values in between small value and big value of learning rate. The combination of various learning rate values will prove to be helpful in increasing the convergence speed of BPNN algorithm and learning accuracy of BPNN algorithm.

2. Literature Review of Back Propagation Algorithm

Backpropagation algorithm is one the most commonly used algorithm of artificial neural network [10,11,12,13]. Backpropagation is applied to solve variety of problems [14,15,16]. The algorithm that is used in this research is the Backpropagation algorithm. This algorithm is considered to be the most suitable and efficient algorithm for the multilayer architectures. In case of presenting unusual pair, a small learning rate is required to use that it can avoid any major disruption [17,18]. Ojhaet. al have applied BPNN algorithm for Detection of proportion of different gas components present in manhole gas mixture [19]. They have used sensor array madeup of semiconductor to detect gas components. They have achieved SSE that is below 0.005 in 450 iterations. Dai and Liu have introduced the concept of competitive learning in BPNN algorithm for classification data [9]. They have used the concept of bucket of weight matrices in their research work. Then weight matrices buckets are used in the competition to select the optimal set of weights. The extra memory used in the bucket cost more resource utilization in this research work. BPNN algorithm is used by Nagathan, Manimozhi andMungara [20] for content based image recognition system. CBIR system fetches similar types of images from any database of images. Image features are given as input to train neural network for this application. The research result shows the 88% precision and 78% recall of image retrieval system. They have used various categories of images like food, buildings, beach, elephants, buses etc in their research work. Khan, Alin and Hussain [21] have used BPNN algorithm for Price Prediction of Share Market. They have used BPNN to predict the forecast stock market. Although the prediction results of BPNN are closed to actual results but few datasets are used to conduct the experimentation. A three term backpropagation algorithm is proposed by zweiri et al [4]. They have used learning rate, momentum term and proportional factor. They have used proportional factor for performance improvement of BP algorithm. Goyal and Goyal have introduced Cascade BPNN algorithm [22] for shelf life detection of processed cheese application. They have conducted their research by observing the mean square error and other assessment parameters of BPNN algorithm for this application. They have used actual sensory score (ASS) and BPNN predicted sensory score (PSS) in their research work. Rubio, Angelov and Pacheco [23] have used uniform backpropoagation algorithm in their research work. They have used a uniform stability theorem for discrete time system. The proposed variation is helpful in online identification and small zone convergence.

Sapna, Tamilarasi and Kumar have introduced Backpropagation Learning Algorithm Based OnLevenberg Marquardt Algorithm [24]. They used BPNN for predicting diabetes disease on the data collected from expert persons and patients. Borucki, Boczar and Cichoń [25] have introduced resilient Backpropagation algorithm for signal recognition. The research result shows that the resilientBackpropagation algorithm is helpful in recognizing the signal adopted neuron classifier at the level exceeding 90%. Reynaldi, Lukas and Margaretha [26] have introduced finite element based neural network. They have used BPNN for differential equation and inverse problem of differential equation. BPNN algorithm successfully solves inverse matrix calculation for solving both differential equation and inverse differential problem. Audhkhasi, Osoba and Kosko [27] have introduced the concept of noise in BP algorithm for convergence speed improvement. They have added the noise in training data of BP algorithm. They have used MNIST digit classification application to assess the performance of BP algorithm. Koscak, Jaksa, and Sincak [28] have used BPNN algorithm for prediction of temperature daily profile. They have used stochastic weight update mechanism in their research work. They used the memory of previous stored step of NN and then used it for prediction. Benzer and Benzer [29] have used BPNN algorithm for prediction of freshwater fish caught in Turkey using data of past 10 years. They have used sum squared error assess performance parameter in their research. The research results are quite efficient for the considered application. Saxena, Jain and Singhal [30] have used BPNN algorithm for hand gesture recognition. They have used android device to recognizing 40 basic hand gestures. They have used centroid of hand, thumb presence and number of peaks in the hand gesture as features in this application. Arulselvi [31] have used BPNN algorithm for Reducing Mismatches in the analog signal processing. BPNN algorithm is helpful to minimize error in Analog VLSI Signals.Yalcin, N., TEZEL, G., &Karakuzu have used hybrid of BPNN algorithm with Particle Swarm Optimization algorithm in their research work [32]. They have used BPNN hybrid for diagnosis of epilepsy disease. Electroencephalogram records are used in training the the hybrid of BPNN algorithm for epilepsy diagnosis.

3. The Back Propagation Algorithm and Proposed Enhancement

This algorithm is one of most normally used learning algorithm and useful to a variety of problems. Simple back propagation is one of the most successful ANN so far invented. BP algorithm has been applied to solve variety of problems. The steps used to train feed forward ANN by BP algorithm are as follows [1]:

Step0: In this step all the weights are initialized to zero or minute random value. Initialize weights.

Step1: This step repeats other steps from step 2 to step 9, till a specific condition remains true.

Step2: In this step the training patterns are used to perform calculation from step 3 to step 8.

Step3: In this step each input neuron receives input from unit (Xi, i=1…..n).

Step4: In this step each hidden unit(Zj, j=1…….n) receives weighted input from input layer and sums it, apply activation function, generates output of the hidden layer. The formula used to sum weight input signals is

$$z\_in_j = v_{0j} + \sum_{i=1}^{n} x_i v_{ij} \quad \text{……… .. (1)}$$

and the formula used to find output signals of hidden layer is Z_j=f(z_in$_j$) where f is the activation function.

Step5: In this step each output unit sums its weighted input signals, and output signal is obtained by applying activation function on it.

$$y\_in_k = w_{0k} + \sum_{j=1}^{p} z_j w_{jk} \quad \text{………(2)}$$

and the formula used to find output signals of hidden layer is y_k=f(y_in$_k$).

Step6: In this step each output unit receives a target pattern corresponding to the input training pattern. The formula is

$$\delta_k = (t_k - y_k) f'(y\_in_k) \quad \text{..........(3)}$$

and then computes its error information terms. Its formula is

$$\Delta w_{jk} = \alpha \delta_k Z_j \quad \text{……………………..........(4)}$$

Step7: The error at hidden layer is calculated by using the formula

$$\delta\_in_j = \sum_{k=1}^{m} \delta_k w_{jk} \quad \text{…..……………(5)}$$

calculates its weight correction term and velocity update using following equations

$$\delta_j = \delta\_in_j f'(z\_in_j) \quad \text{………….……(6)}$$

$$\Delta v_{ij} = \alpha \delta_j x_i \quad \text{…… ………………… ………(7)}$$

Step8: Update the weights from hidden layer to output layer and weights from input layer to hidden layer. The formula is as below

$$w_{ij}(new) = w_{jk}(old) + \Delta w_{jk} \quad \text{……………....…(8)}$$

$$v_{ij}(new) = v_{ij}(old) + \Delta v_{jk} \quad \text{….……..….(9)}$$

Step9: In this step the stopping condition is tested.

This research work introduces five variable learning rate schemes in BPNN algorithm. These variable learning rates will take various values of learning rate from a specific interval of learning rate. The learning rate schemes include learning mechanism of increasing manner, decreasing manner learning rate, the chaotic manner and oscillating manner of variable change in the learning rate. A hybrid learning rate scheme is also included in this research that consist of a randomized pool of other five learning rate have different behaviors to check the behavior of BPNN algorithm over the fusion of various nature leaning rates. Learning mechanism of increasing manner, decreasing manner learning rate, the chaotic manner and oscillating manner of variable change in the learning rate are proposed in this paper. A hybrid learning rate scheme is also included in this research that consist of a randomized pool of other five learning rate have different behaviors to check the behavior of BPNN algorithm over the fusion of various nature learning rates.

3.1 Linear Increasing learning Rate

In this method we start from a small value of learning rate and then linearly increases that small value to a large value. The small value is taken to 0.1 and large value is taken to 0.9 The Linear increasing inertia weight mathematical equation is as,

$$alpha_i = alpha_{\min} + (alpha_{\max} - alpha_{\min}) \times \frac{Max\_iter - current\_iter}{Max\_iter} \quad \ldots\ldots\ldots\ldots(10)$$

Where $alpha_{\max}$ is maximum value of alpha and $alpha_{\min}$ is the minimum value of alpha. The value of $alpha_{\max}$ is taken to 0.9 and value of $alpha_{\min}$ is taken as 0.1. Max_iter is the maximum number of training iterations used to train BPNN algorithm and current_iter is current iteration for which alpha is being calculated.

3.2 Linear Decreasing learning Rate

In linearly decreasing inertia weight, large value linearly decreased to small value. The small value is taken to 0.1 and large value is taken to 0.9 Linear decreasing learning rate is following.

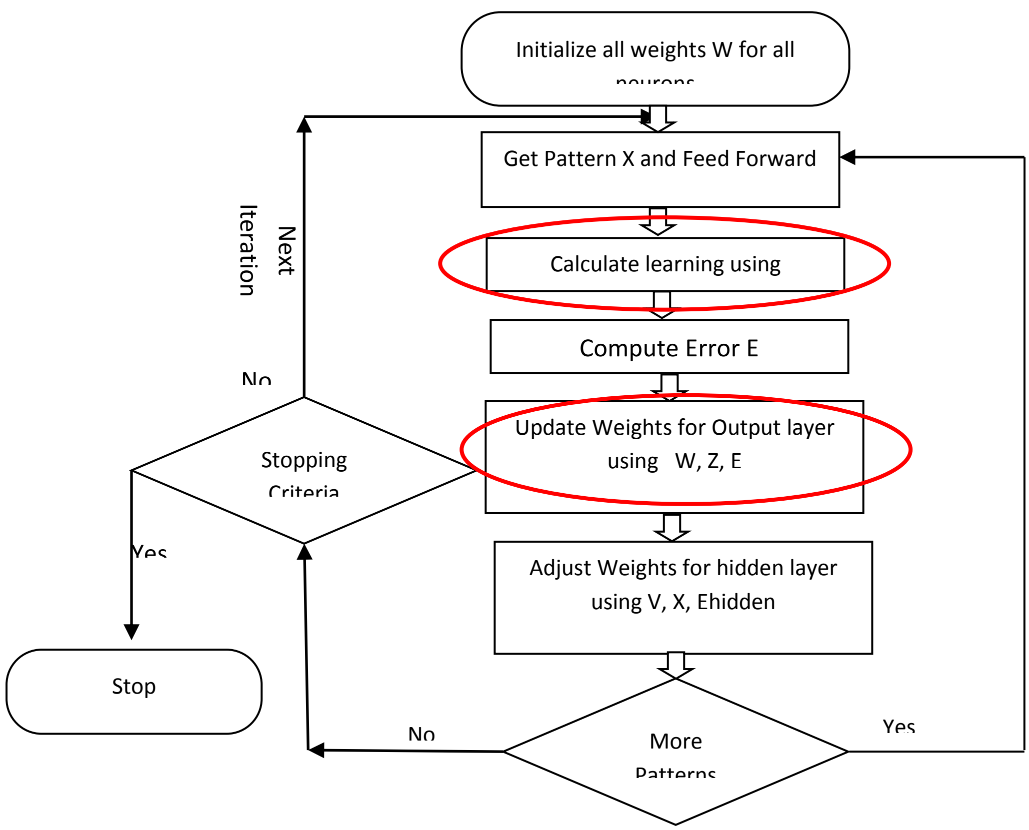
$$alpha_i = alpha_{\max} - (alpha_{\max} - alpha_{\min}) \times \frac{Max\_iter - current\_iter}{Max\_iter} \quad (11)$$

Where $alpha_{\max}$ is maximum value of alpha and $alpha_{\min}$ is the minimum value of alpha. The value of $alpha_{\max}$ is taken to 0.9 and value of $alpha_{\min}$ is taken as 0.1. Max_iter is the maximum number of training iterations used to train BPNN algorithm and current_iter is current iteration for which alpha is being calculated.

3.3 The Chaotic learning Rate

This strategy was introduced by "chaotic optimization mechanism" is successfully applied in PSO to choose the value of inertia weight. In this case chaotic mapping is used for learning rate that is used in the training of BPNN algorithm. The equation of chaotic learning rate is as under where random number z is selected between interval 0 and 1.

$$alpha_i = \frac{(alpha_{\max} - alpha_{\min}) \times Max\_iter - current\_iter}{Max\_iter + alpha_{\min} \times z} \quad (12)$$



3.4 Random Chaotic learning Rate

In this strategy a random number z belongs to (0 , 1) is generated. In addition to z; a random number Rand()is also generated from the range (0, 1). The formula of random Choatic learning rate is.

$$alpha_i = 0.5 \times Rand() - 0.5 \times z \quad \text{............} \quad .(13)$$

3.5 Oscillating Learning Rate

In this kind of inertia weight strategy, a wave of global search is following by a wave of local search. This technique is implemented for learning rate using the following equation

$$T = \frac{2S}{3+2k} \quad (14)$$

$alpha_{\min}$ is the starting learning rate and $alpha_{\max}$ is the

$$alpha_i = \frac{(alpha_{\max} + alpha_{\min})}{2} + \frac{(alpha_{\max} - alpha_{\min})}{2} * \cos\left(\frac{2\pi t}{T}\right)$$

ending learning rate. S is number if training iterations used to train BPNN algorithm while k is used to control the frequency of oscillations

3.6 Hybrid Learning Rate

In this learning rate method we have used a pool of above five learning rates. During each training iteration of BPNN algorithm, we have used one learning rate method. The learning rate in hybrid method is selected from the pool in simple random manner. That pool contains learning rate of different nature that will be helpful to incorporate diversity in leaning rate and learning process of BPNN algorithm. The random number generates a random number in the range [1,5] to select the strategy from the pool to be used during the current iteration.

4. Experimental setup and Classification Applications

To evaluate the performance of the simple Backpropagation and proposed variations are analyzed based on the mean square error and accuracy of results for all classification problems. For fair comparison the number of hidden neurons in the hidden layer and number of training iterations are used same for all classification applications. The number of hidden neurons are taken 12 and 18 for all applications; total number of training iterations used are taken as 500 and log sigmoid activation function is used at output layer. The training and testing has done on separate date. The data that was included in the training was not included in the testing. In testing the unseen data is presented to get the simulation results. The 15% of the data is used for training and 85% of the data is used in testing. The 15% data that is used in testing in not included in the training data and equal number of instances are used for each class.

The applications that are considered in this research are the Iris classification problem, Breast Cancer problem and wine classification taken from the Standard UCI datasets. The detail of each application is given as follows

a) Iris Classification problem

This is a very important problem of classification used by different researchers. It is one of the most used classification class of iris plant dataset from the UCI datasets. The data in this problem is taken from the UCI standard datasets. The detail of attributes of this problem is as follows. This data set has four attributes; Sepal length, Sepal width, petal length and petal width. Iris problem has three classes; Iris Setosa, Iris Versicolour and Iris Virginica. The total numbers of attributes that are used in this application are 150. That 150 contain 33.3% data of each of the three classes.

b) Breast Cancer Problem

This is a very important problem of computing and it is widely used classification problem used by different researchers. The data in this problem is taken from the UCI standard datasets. The lists of attributes that are used in breast cancer classification are as follows. These nine attributes are Age, Menopause, Tumor-size, Inv-nodes, Node-caps, Deg-malig, Breast, Breast-quad and Irradiate. The numbers of instances that are used in this class are the 201 and 85 instances respectively for recurrent and non-recurrent classes.

c) Wine Classification Problem

This is a very important problem of computing and it is widely used classification problem used by different researchers. The data in this problem is taken from the UCI standard datasets. The attributes that are used in wince classification includes Alcohol, Malic acid, Ash, Alcalinity of ash, Magnesium, Total phenols, Flavanoids, Nonflavanoid phenols, Proanthocyanins, Color intensity, Hue, OD280/OD315 of diluted wines and Proline

.

Table. 1 Standard BPNN, LIBPNN, LDBPNN, OBPNN, TCBPNN, CBPNN and HBPNN Accuracy and Mean Square Error for Iris Data

Algorithm	Performance parameter	12 Hidden Neurons	18 Hidden Neurons
BPNN	Accuracy	73.33%	80.00%
	MSE	0.008211	0.00881
LIBPNN	Accuracy	73.33%	66.67%
	MSE	0.003553	0.003424
LDBPNN	Accuracy	80.00%	100.00%
	MSE	0.007385	0.007056
OBPNN	Accuracy	66.67%	86.67%
	MSE	0.004877	0.006437
TCBPNN	Accuracy	86.67%	100.00%
	MSE	0.004088	0.00506

CBPNN	Accuracy	86.67%	40.00%
	MSE	0.006261	0.006247
HBPNN	Accuracy	73.33%	93.33%
	MSE	0.004694	0.00313

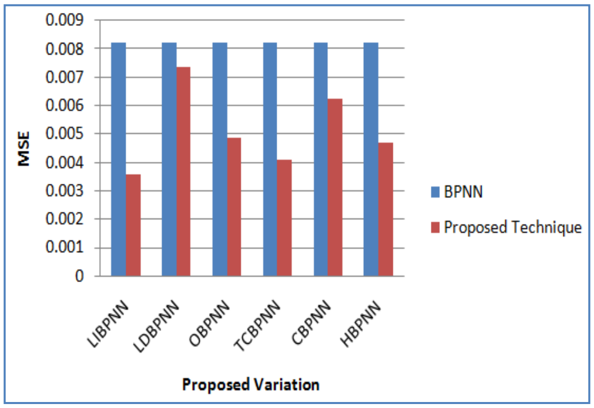


Figure.3 Mean Square Error of BP algorithm and proposed variations using 12 hidden neurons for Iris Data

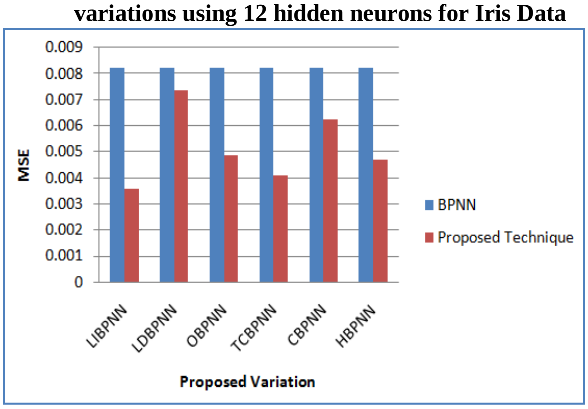


Figure 3 and Figure 4 contains MSE of BP algorithm and proposed variation for IRIS dataset. It is clear from Figure 3

Figure. 4 Mean Square Error of BP algorithm and pro posed variations using 18 hidden neurons for Iris Data for IRIS

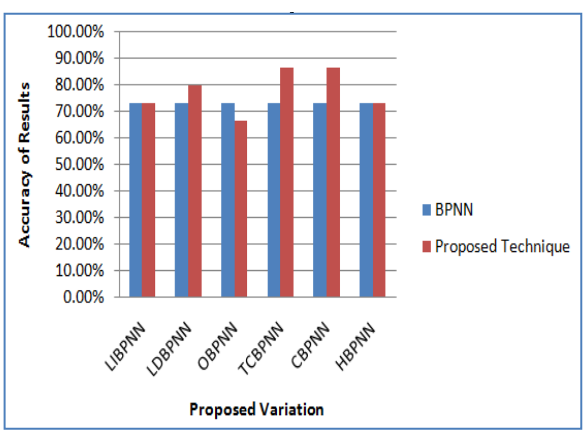


Figure. 5 Accuracy of Results of BP algorithm and proposed variations using 12 hidden neurons for Iris D

and Figure 4 that MSE of LIBPNN, LDBPNN, OBPNN, TCBPNN, CBPNN, HBPNN is better than BPNN in all cases of 12 hidden neurons and 18 hidden neurons.

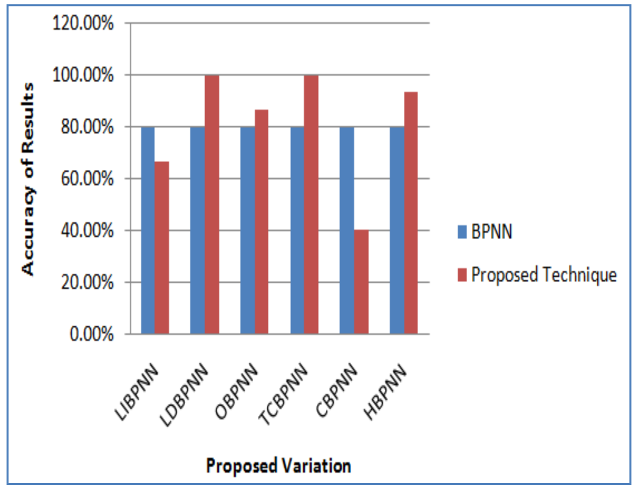


Figure. 6 Accuracy of Results of BP algorithm and proposed variations using 18 hidden neurons for Iris Data

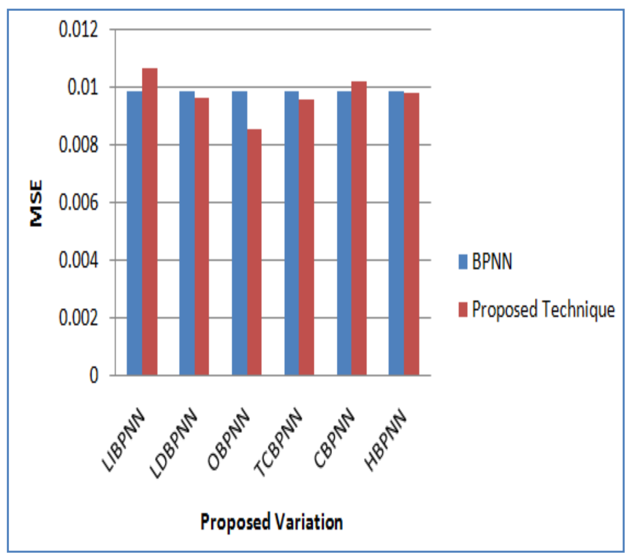


Figure. 7 Accuracy of Results of BP algorithm and proposed variations using 12 hidden neurons Breast

5. SIMULATION RESULTS

The simulation results are obtained for standard backpropagation and proposed variations in BP algorithm. The results of standard backpropagation and proposed variations are taken by taking same number of hidden neurons in the hidden layer and fixed training iterations throughout the experimentation. The results of standard backpropagation algorithm are taken as the average of results obtained using five different values of learning rate where alpha={0.1, 0.3, 0.5, 0.7, 0.9} throughout the experimentation. The fix training iterations that are equal to 500 are used to get the experimental results. The numbers of hidden neurons in the hidden layer are used as 12 and 18 in the experimentation. The experimental results contains Testing Accuracy in tabular form, Mean Square Error (MSE) in tabular form, MSE graphical presentation and Accuracy of Results in graphical form

5.1 Iris Classification Results

The simulation results of standard Backpropagation algorithm (BPNN) and other proposed enhancements (LIBPNN, LDBPNN, OPBNN, TCBPNN, CBPNN and HBPNN) results for Iris classification problem are given in this section. The accuracy and MSE of iris data are given in table 1. MSE results in graphical form are given in Figure 3 and 4 while Accuracy results are given in Figure 5 and 6. Figure 5 and Figure 6 contains testing accuracy of BP algorithm and other proposed variation. From Figure 5 and Figure 6 contains accuracy of testing results of BP algorithm and proposed variation for IRIS dataset. From Figure 5 and Figure 6 it is clear that testing accuracy of LDBPNN, TCBPNN, CBPNN is better, HBPNN, LIPNN is similar and OBPNN is worse for 12 hidden neurons. The performance of LDBPNN, POBPNN, TCBPNN and BBPNN is better while LIBPNN, CBPNN is worse than BPNN in case of 18 hidden neurons.

5.2 Breast Cancer Results

The simulation results of standard Backpropagation algorithm (BPNN) and other proposed enhancements (LIBPNN, LDBPNN, OPBNN, TCBPNN, CBPNN and HBPNN) results for Breast Cancer classification problem are given in this section. The accuracy and MSE of iris data are given in table 2. MSE results in graphical form are given in Figure 7 and 8 while Accuracy results are given in Figure 9 and 10.

Table.2 Standard BPNN, LIBPNN, LDBPNN, OBPNN, TCBPNN, CBPNN and HBPNN Accuracy and Mean Square Error for Breast Cancer Data

Algorithm	Performance parameter	12 Hidden Neurons	18 Hidden Neurons
BPNN	Accuracy	91.43%	86.57%
	MSE	0.00985	0.009531
LIBPNN	Accuracy	92.14%	87.14%
	MSE	0.010697	0.010723
LDBPNN	Accuracy	92.86%	87.86%
	MSE	0.009639	0.009509
OBPNN	Accuracy	92.86%	95.71%
	MSE	0.008557	0.006871
TCBPNN	Accuracy	94.29%	87.14%
	MSE	0.009595	0.009988
CBPNN	Accuracy	93.57%	92.86%
	MSE	0.010193	0.009231
HBPNN	Accuracy	91.43%	97.86%
	MSE	0.009834	0.009631

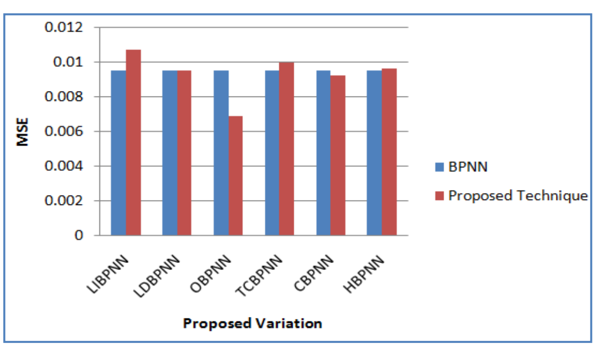


Figure. 8 Accuracy of Results of BP algorithm and proposed varitions using 18 hidden neurons Breast

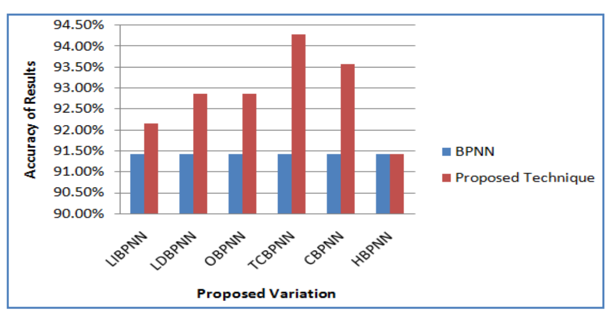


Figure. 9 Accuracy of Results of BP algorithm and proposed variations using 12 hidden neurons Breast

Figure 7 and Figure 8 contain MSE of BP algorithm and proposed variation for the breast cancer dataset. From Figure 7 and Figure 8 it is clear that MSE of LDBPNN, OBPNN, TCBPNN, HBPNN is better than BPNN for 12 hidden neurons, while LIBPNN and CBPNN is worse than BPNN for 12 hidden neurons. In case of 18 hidden neurons the performance of LDBPNN, OBPNN, CBPNN and HBPNN is better while LIBPNN, TCBPNN is worse than BPNN algorithm.

Figure 9 and Figure 10 contains testing accuracy of BP algorithm and other proposed variation for Breast Cancer dataset. From Figure 9 and Figure 10 it can be observed that for testing accuracy of LIBPNN, LDBPNN, OBPNN, TCBPNN, CBPNN and HBPNN is better than BPNN in all cases except one case where performance of HBPNN is equal to performance of BPNN for 18 hidden neurons.

5.3 Wine Classification Results

.

Table 5.1 Standard BPNN, LIBPNN, LDBPNN, OBPNN, TCBPNN, CBPNN and HBPNN Accuracy and Mean Square Error for Wine Data

Algorithm	Performance parameter	12 Hidden Neurons	18 Hidden Neurons
BPNN	Accuracy	72.78%	62.78%
	MSE	0.010252	0.000461
LIBPNN	Accuracy	88.89%	50.00%
	MSE	0.011604	0.006151
LDBPNN	Accuracy	80.56%	50.00%
	MSE	0.00382	0.000327
OBPNN	Accuracy	77.78%	50.00%
	MSE	0.004559	0.000136
TCBPNN	Accuracy	61.11%	72.22%
	MSE	0.006843	0.000265
CBPNN	Accuracy	75.00%	61.11%
	MSE	0.009964	0.000389
HBPNN	Accuracy	80.56%	86.11%
	MSE	0.00452	0.00021

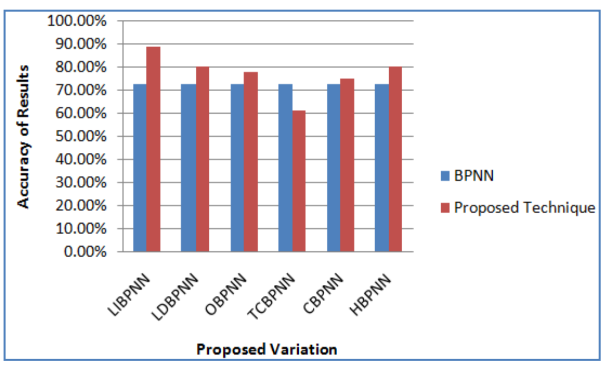


Figure. 10 Accuracy of Results of BP algorithm and proposed variations using 18 hidden neurons Breast Cancer Data

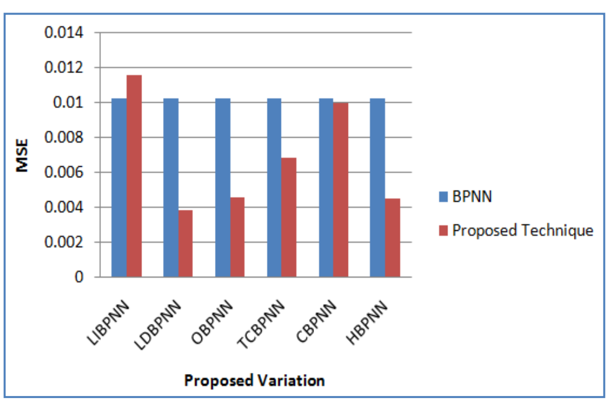


Figure. 11 Mean Square Error of BP algorithm and proposed variations using 12 hidden neurons for Wine Data

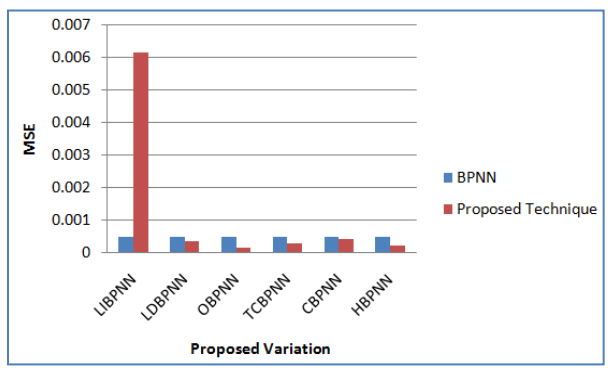


Figure. 12 Mean Square Error of BP algorithm and proposed variations using 18 hidden neurons for Wine Data

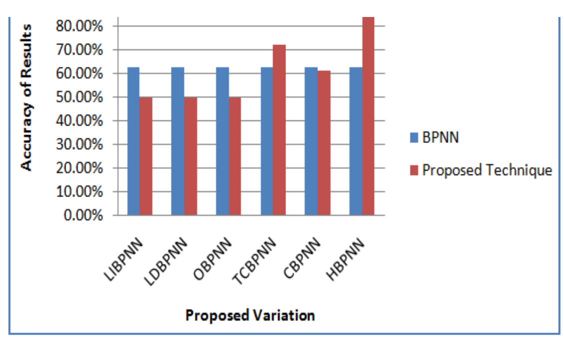


Figure. 13 Accuracy of Results of BP algorithm and proposed variations using 12 hidden neurons for Wine

The simulation results of standard Backpropagation algorithm (BPNN) and other proposed enhancements (LIBPNN, LDBPNN, OPBNN, TCBPNN, CBPNN and HBPNN) results for Wine classification problem are given in this section. The accuracy and MSE of iris data are given in table 2. MSE results in graphical form are given in Figure 7 and 8 while Accuracy results are given in Figure 9 and 10

Figure 11 and Figure 12 contains MSE of BP algorithm and proposed variation for Wine Classification dataset. From Figure 11 and Figure 12 it is clear that MSE of all proposed variations except LIBPNN is better than BPNN in both cases.

Figure 13 and Figure 14 contains testing accuracy of BP algorithm and proposed variation for Wine Classification dataset. From Figure 13 and Figure 14 it is clear that performance of all proposed variations is better than BPNN except TCBPNN in case of 12 hidden neurons while performance of TCBPNN and HBPNN is better than BPNN for 18 hidden neurons.

The performance of proposed variations in learning rate of BP algorithm are compared with standard BP algorithm over commonly used standard UCI machine learning repository datasets. For fair comparison two cases of number of hidden neurons (case-1,12 nodes and case-II contains 18 nodes) are kept constant throughout the experimentation. The training iterations are also taken constant to 500. The results of testing accuracy and mean square error are obtained during the simulation process. The convergence graphs of mean square error are also obtained during the simulation process that are presented in graphical for each standard dataset and each technique including linear increasing Backpropagation neural network (LIBPNN), linear decreasing Backpropagation neural network (LDBPNN), Oscillating Backpropagation neural network (OBPNN), The Chaotic Backpropagation neural network (TCBPNN), Chaotic Backpropagation neural network (CBPNN) and hybrid of all these techniques that is HBPNN are better than BPNN in most of the cases and comparable in some cases. The proposed learning rate technique has better performance in most of the cases and comparable performance in some cases of error convergence graphs. The results are presented in tabular form as well as graphical form. The accuracy of results given in tabular as well as graphical form shows that LIBPNN, LDBPNN, OBPNN, TCBPNN, CBPNN and HBPNN is better than BPNN in most of the cases, and performance is comparable in some cases and performance of BPNN is better in few cases. From experimental results it can be summarized that experimental results of proposed variations in most of the cases are better or comparable to BP algorithm. The experimental results of linear increasing Backpropagation neural network (LIBPNN), linear decreasing Backpropagation neural network (LDBPNN), Oscillating Backpropagation neural network (OBPNN), The Chaotic Backpropagation neural network (TCBPNN), Chaotic Backpropagation neural network (CBPNN) and hybrid of all these techniques that is HBPNN are better than in most of the cases and comparable in some cases.

CONCLUSION

Learning rate has a key role in Weight updation using Backpropagation algorithm of ANN. This thesis introduces some new learning schemes that include linear increasing learning rate, linear decreasing learning rate, oscillating learning rate, The Chaotic learning rate, Chaotic learning rate and a hybrid learning rate that is formed by creating a randomized pool of these five learning rate schemes. These learning rate schemes have characteristics of different nature like simple increasing, simple decreasing, chaotic manner, oscillating manner etc. Weights in ANN are key learning factor that learns from the training patterns. The updation of weight includes old weight values, training patter, network error and target value. The value of learning rate affects the performance of BPNN algorithm. The small value of learning rate slows down the learning mechanism and big value of learning rate over trains ANN architecture. The new learning rate schemes includes small, big and values in between interval of small and big values. The various nature of learning rate is analyzed that will be helpful in using variable learning rate instead of constant learning rate. The performance metric that is used in the analysis of results includes two very common performance parameters; the testing accuracy and the mean square error. Based on these performance parameters research results of variable learning rate schemes are compared with the constant learning rate schemes that show significance performance over the constant learning rate. The constant learning rate results are obtained by taking the mean of {0.1, 0.3, 0.5, 0.7, 0.9} values results. These results are compared with the proposed variable learning rates. Research results shows that proposed learning rate schemes have dominating performance in most of the cases. The future work of this work can be to analyze and adapt pool of learning rates that will be helpful in quick convergence of BPNN algorithm. The adaptive and self adaptive learning based on the learning can prove to be a valuable addition in BPNN algorithm. Another possible direction of future work is to implement the proposed variable learning rate based techniques in deep learning neural networks.

Conflict of interest statement: The authors declare that there is no conflict of interests regarding the publication of this paper.

REFERENCES

[1] L. Fausett, *Fundamentals of neural networks: architectures, algorithms, and applications*, 1st ed. NJ, USA: Prentice-Hall, Inc., 1994.

[2] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. NJ, USA: Prentice Hall, 1994.

[3] M. A. Arbib, *The Handbook of Brain Theory and Neural Networks*, 2nd ed. USA: MIT Press Cambridge, 2002.

[4] Y. Zweiri, J. F. Whidborne, and L. D. Seneviratne, "A three-term backpropagation algorithm," *Neurocomputing*, vol. 50, pp. 305-318, 2003.

[5] X. Wang and al. et, "An improved backpropagation algorithm to avoid the local minima problem," *Neurocomputing*, vol. 56, pp. 455-460, 2004.

[6] J. Wei and et. al, "Subgrade settlement prediction of transition section based on momentum back-propagation," *Journal of Beijing Jiaotong University*, vol. 1, no. 11, pp. 12-18, 2012.

[7] H. H. Orkcu and H. Bal, "Comparing performances of backpropagation and genetic algorithms in the data

classification," *Expert Systems with Applications*, vol. 38, no. 4, pp. 3703–3709, 2011.

[8] W. Du, "Research on Evaluation of Enterprise Performance Based on BP Neural Network Improved by Levenberg-marquardt Algorithm.," in *International Conference on Automation, Mechanical Control and Computational Engineering*, 2015.

[9] Q. Dai and N. Liu, "Alleviating the problem of local minima in Backpropagation through competitive learning," *Neurocomputing*, vol. 94, no. 1, pp. 152–158, 2012.

[10] K. Sridevia, E. Sivaramanb, and P. Mullaia, "Back propagation neural network modelling of biodegradation and fermentative biohydrogen production using distillery wastewater in a hybrid upflow anaerobic sludge blanket reactor," *Bioresource Technology*, vol. 165, pp. 233-240, 2014.

[11] B. T. Nukala and et. al, "An Efficient and Robust Fall Detection System Using Wireless Gait Analysis Sensor with Artificial Neural Network (ANN) and Support Vector Machine (SVM) Algorithms," *Open Journal of Applied Biosensor*, vol. 3, no. 4, pp. 29-35, 2015.

[12] P. Karthigayani and S. Sridhar, "Decision tree based occlusion detection in face recognition and estimation of human age using back propagation neural network," *Journal of Computer Science*, vol. 10, no. 1, pp. 115-121, 2014.

[13] J. S. Kim and S. Jung, "Implementation of the RBF neural chip with the back-propagation algorithm for on-line learning," *Applied Soft Computing*, vol. 29, pp. 233–244, 2015.

[14] N. Lopes and B. Ribeiro, "Towards adaptive learning with improved convergence of deep belief networks on graphics processing units," *Pattern Recognition*, vol. 47, no. 1, pp. 114-127, January 2014.

[15] S. Taghadomi-Saberi and et. al, "Evaluating the potential of artificial neural network and neuro-fuzzy techniques for estimating antioxidant activity and anthocyanin content of sweet cherry during ripening by using image processing," *Journal of the Science of Food and Agriculture*, vol. 94, no. 1, pp. 95-101, 2014.

[16] N. N. Nawi, M. Z. Rehman, and A. Khan, "A new bat based back-propagation (BAT-BP) algorithm," *Advances in Systems Science*, pp. 395-404, 2014.

[17] D. Ciresan and et. al, "Multi-column deep neural networks for image classification," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3642-3649.

[18] N. Srivastava and et. al, "Dropout: A simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929-1958, 2014.

[19] V. K. Ojha and et. al, "Detection of proportion of different gas components present in manhole gas mixture using backpropagation neural network," in *International Conference on Information & Network Technology*, 2012, pp. 20-26.

[20] A. Nagathan, Manimozhi, and J. Mungara, "Content-Based Image Retrieval System using Feed-Forward Backpropagation Neural Network," *International Journal of Computer Science and Network Security*, vol. 14, no. 6, pp. 70-77, 2014.

[21] Z. H. Khan, T. S. Alin, and M. A. Hussain, "Price prediction of share market using artificial neural network (ANN)," *International Journal of Computer Applications*, vol. 22, no. 2, pp. 42-47, 2011.

[22] S. Goyal and G. K. Goyal, "A novel method for shelf life detection of processed cheese using cascade single and multi layer artificial neural network computing models," *ARPN Journal of Systems and Software*, vol. 2, no. 2, pp. 79-83, 2012.

[23] J. D. Rubio, P. Angelov, and J. Pacheco, "Uniformly stable backpropagation algorithm to train a feedforward neural network," *IEEE Transactions on Neural Networks*, vol. 22, no. 3, pp. 356-366, 2011.

[24] S. Sapna, A. Tamilarasi, and M. P. Kumar, "Backpropagation Learning Algorithm Based On Levenberg Marquardt Algorithm," *omputer Science & Information Technology*, pp. 393-398, 2012.

[25] S. Borucki, T. Boczar, and A. Cichoń, "The application of the resilent backpropagation algorithm and power spectrum density for recognizing the acoustic emission signals genereted by basic partial discharge forms using artificial neuron networks," *Archives of Acoustics*, pp. 335-342, 2014.

[26] A. Reynaldi, S. Lukas, and H. Margaretha, "Backpropagation and Levenberg-Marquardt algorithm for training finite element neural network," in *6th European Symposium on Computer Modeling and Simulation*, 2012, pp. 89-94.

[27] K. Audhkhasi, O. Osoba, and B. Kosko, "Noise benefits in backpropagation and deep bidirectional pre-training.," in *The 2013 International Joint Conference on Neural Networks*, 2013, pp. 1-8.

[28] J. Koscak, R. Jaksa, and P. Sincak, "Prediction of temperature daily profile by stochastic update of backpropagation through time algorithm," in *International Symposium on Forecasting*, 2011, pp. 1-9.

[29] R. Benzer and S. Benzer, "Application of artificial neural network into the freshwater fish caught in Turkey," *International Journal of Fisheries and Aquatic Studies*, vol. 2, no. 5, pp. 341-346, 2015.

[30] A. Saxena, D. K. Jain, and A. Singhal, "Hand Gesture Recognition using an Android Device," in *Fourth IEEE International Conference on communication Systems and Network Technologies*, 2014, pp. 819-822.

[31] B. B. Arulselvi, "Reducing Mismatches in the Analog Signal by Using Levenberg-Marquardt Back Propagation Algorithm," *Middle-East Journal of Scientific Research*, vol. 20, no. 12, pp. 2346-2353, 2014.

[32] N. Yalcin, G. Tezel, and C. Karakuzu, "Epilepsy diagnosis using artificial neural network learned by PSO," *urkish Journal of Electrical Engineering and Computer Science*, vol. 23, no. 2, pp. 421-432, 2015.

[33] J. T. Huang and et. al, "Cross-language knowledge transfer using multilingual deep neural network with shared hidden layers," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 7304-7308.