APPLICATION AND ANALYSIS OF ARTIFICIAL NEURAL NETWORK BACKPROPAGATION ALGORITHM'S IN KNOWLEDGE MANAGEMENT

Shailendra Singh[1], Venkataramanaiah Saddikuti[1], Shantanu Bhattacharya[2]

[1]Indian Institute of Management, Lucknow, India

[2]Department of Mechanical Engineering, Indian Institute of Technology, Kanpur, India

Abstract

Knowledge transfer impacts the performance of any firm and its adaptability to the environment and associated change. Various knowledge management models capture the process of knowledge transfer and also help in explaining the efficacy of the transfer process. Artificial Neural Networks (ANN) can help to simulate and predict the knowledge transfer outcomes in the same way as the human brain perceives and interprets information to provide any learning outcome. However ANN has various algorithms each of which is an independent direction. This study discusses the direct application of Artificial Neural Network (ANN) framework to execute knowledge transfer using multiple ANN algorithm's and perform efficacy and ability determination to estimate the appropriateness of the knowledge transfer method. Virtual reality based simulated environment is used as a vehicle to make the knowledge transfer and the ability of subjects of various categories to entrain on warehousing management domains is realized. The ANN applies to prediction of the knowledge level of the trained subjects initially by pre-training the ANN through multiple algorithms and then finding the closeness of the outcomes predicted through this network using multiple algorithms to the real data which is gauged through real examination of the trained subjects.

Keywords: *knowledge management, knowledge transfer, artificial neural network, backpropagation, fletcher-reeves updates, polak- ribiére updates, ann*

1. INTRODUCTION

Knowledge creation [1], Knowledge management [2] and knowledge transfer [3] are very important areas of research and there have been organizations who have directly applied organized studies in these domains to increase the key performance indices [4] and drive to a motivated and go getter state, In context of the same growing technology areas like virtual systems [5], virtual world [6], augmented reality training platforms [7], Artificial Neural Networks [8], Simulated environment [9] etc. have been heavily deployed to make modern day knowledge systems and transfer models. Of significant mention is the use of self-training networks for predictive outcomes related to knowledge models and various ANN algorithms like the "Conjugate gradient backpropagation algorithm with Fletcher-Reeves updates" [10], or that with Polak- Ribiére updates [11] can turn the predictability of system into various accuracy levels .

Study done at various educational institutes and lead industries show that the way forward to transfer knowledge from amongst various vehicles like hands on lab, animated videos, lab video and lecture mode is animated videos to maximize the transfer rate (Singh et al, 2020). Artificial neural networks (ANN) are widely used in different domains like medical science (Trappey et al 2013), electrical systems (A Singh et al, 2014), Knowledge Warehousing (Nemati et al, 2002) etc. There are various knowledge models proposed by different thought leaders such as the SECI model (Nonaka et al , 2000), sense making model [13], Adaptive knowledge transfer model (14) etc. Knowledge transfer can influence any firm's performance in terms of ability to meet the KPI targets [20] In this paper we attempt to predict the outcome of knowledge transfer through a layer of neurons in an ANN and apply various algorithms like "Conjugate gradient backpropagation with Fletcher-Reeves updates and Polak-Ribiére updates to perform a comparative of the outcomes predicted through both algorithms. It is found that predictions of the Polak- Ribiére algorithm is closer to the reality plots generated through response examination of trained subjects in real life,

2. THEORY

2.1 Conjugate gradient Backpropagation with Fletcher-Reeves updates (Fletcher, Reeves, 1964)

The syntax for MATLAB is “traincgf” which is training function for network that updates weight and bias values as per the “conjugate gradient back-propagation algorithm” assisted through “Fletcher-Reeves updates” [10,12]

[net,TR] = traincgf (net,TR,trainV,valV,testV).

The Algorithm “traincgf” can train any network which has involvement of the derivative function in its weight, net input, and transfer functions. It calculates derivatives of performance with respect to the weight and bias variables ‘X’ using Backpropagation. Each variable is adjusted according to the following:

X = X + a*dX; where ‘dX’ is the search direction. In order to minimize the performance along the search direction the parameter ‘a’ is selected. The minimum point is located using the line search function‘search Fcn’. The first search direction is the negative of the gradient of performance. Therafter in following iterations the search direction is computed from the new gradient and the previous search direction, and we use formula i.e.

dX = -gX + dX_old*Z;

where ‘gX’ is the gradient which is the gradient of iterative X. The parameter Z can be computed in various ways. For the “Fletcher-Reeves variation of conjugate gradient” it is computed according to

Z= [normnew_sqr/norm_sqr]; where [norm_sqr] is the norm square of the previous gradient and normnew_sqr is the norm square of the current gradient.

Training stops on achievement of conditions i.e.

1. When no. of epochs (repetitions) reached are maximum
2. When max time is exceeded.
3. When Performance is minimized to goal.
4. When performance gradient falls below min_grad.
5. When validation performance has increased more than max_fail times from the last time it decreased.

2.2 Conjugate gradient Backpropagation with Polak-Ribiére updates (Polak, Ribiére, 1969)

It uses the function “traincgp” with syntax [net,TR] = traincgp (net,TR,trainV,valV,testV). This is a network training function that updates weight and bias values according to conjugate gradient back-propagation with Polak-Ribiére updates [11,12]and returns trained network

A feed-forward network is created with a hidden layer of 2 neurons. This Algorithm can train any network which has involvement of the derivative function in its weight, net input, and transfer functions

Backpropagation concept is same as explained above in section 2.1

For the “Polak-Ribiére variation of conjugate gradient”, it is computed according to

Z = ((gX - gX_old)'*gX)/norm_sqr; where norm_sqr is the norm square of the previous gradient, and gX_old is the gradient on the previous iteration. Training stops as explained in above sec 2.1

S.No	**Fletcher, Reeves Updates ANN**	**Polak, Ribiére Updates ANN**
1	Z= [normnew_sqr/norm_sqr]	Z = ((gX - gX_old)'*gX)/norm_sqr
2	$\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}.$	$\beta_k^{PR} = \frac{g_{k+1}^T y_k}{g_k^T g_k}.$

3. DETAILED METHOD OF STUDY:

3.1 Survey of the pre-trained and tutorial design for knowledge transfer

In this study we have designed an experimental survey among 450 subjects who have different aspirations and motivation levels and are from related and unrelated backgrounds and further have different interpretation, and learning abilities and also have different skill sets etc. The survey is designed to gauge the trainability index among the participants and then such gauged participants are made to undertake an augmented reality training module with simulated animation as an instructional medium. The augmented reality module is developed using Microsoft Kinect for Windows with field of view 57^o H, 43^o V, Sensors RGB & Depth, Range 1.2m – 3.5m, Depth Steam QVGA (320x240) 16-bit @ 30 fps VGA, RGB Stream VGA (640x480) 32-bit @ 30 fps, Microphones 2 left and 2 right, OS Support Win10 with 64 bit.

3.2 Methods

Pre and post tests have been conducted based on four specific key performance indicators (KPI) viz., Safety, Quality, Productivity and Cost and the outcome of these survey responses have been captured into a ANN model and system level training has been further carried for network learning purposes using the survey responses as inputs and outputs respectively. Once the network is fully trained using the above said data on MATLAB tool using Artificial Neural Network toolbox, another set of 50 subjects have been studied using this trained network through the utilization of the Backpropagation Fletcher Reeves and Polak-Ribiére updates in ANN. The data was earlier reported for the utilization of the LM algorithm which provided a reasonable output although it provided a scope to optimize further the predicted outcome. (Singh et al 2020, submitted to Journal of Knowledge Management, under review)

To begin with a virtual reality animated simulation is prepared as detailed above. The simulated environment is designed to provide experiential training to subjects imbibing in them a feeling of heights etc. while doing realistic work on this VR platform in a real industry environment.

The surveys that are provided to the subjects are designed for two stages that is for the pre-training state and post-training state of the subjects and these are focused to assess four different parameters (called KPIs) viz Safety, Quality, Productivity and Cost. The questions that are asked in the survey are about technical content related to various equipment and processes. These outcomes of these surveys (450 in number) are used to answer before and after the exposure to the virtual experiential training modules. The first round of survey was conducted in this manner.

The outcomes are randomized through randomization of the questionnaire sequence for all the subjects.

3.3 Experimental study

The experimental study is conducted at an industry in India and the results of pre and post test are used to train an ANN model for providing predictive outcomes of the post training performance of the candidates. The experiments designed on the VR platform were related to the height of working operations using equipment which otherwise may be very difficult to learn and also saves accidents which could have happened if physically conducted. The realistic training involves a variety of skill sets due to the processing intricacies etc. It also has a high overall safety risk, affecting the quality of work delivery, and impacts the overall productivity and cost for any warehousing operation to be successfully carried out. Another very important from an industry point of view is the situation created due to pandemic where physical contact is advised against and multiple people working together is avoided if not necessary. The subjects in the training group possessed 10+ years of formal elementary education followed by a specialized technical study for 2+ years. All participants were either semi-skilled or fully skilled with minimum of two years technical education.

Following are the details of the experimental study done:

3.4 Age and experience of the candidates:

The age parameter of the subjects were divided into 4 different groups comprising of subjects whose age would fall between 18-20 years, 20-25 years, 25-30 years, and 30-40 years, respectively. The industry experience of the candidates is also classified into 4 groups with those possessing no experience 0 years, and those having experience of up to 1-year, 1-5 years, and 5-10 years, respectively. However, none of the subjects had any kind of pre-training to operate the target equipment to work at height for which the training module was designed. The table 1 below presents a distribution of no. of subjects lying in a particular group in terms of years of age and that of experience. The total no. of subjects is 450.

Age (years)	**18-20**	**20-25**	**25-30**	**30-40**
Nos	175	202	51	22
Experience (years)	**Nil**	**Upto 1 year**	**1-5 years**	**5-10 years**
Nos	32	73	272	73

Table 1. Age and Experience of Subjects (Total 450 nos)

3.5. Experimental setup

All the 450 participants were assessed before and after the virtual reality animated video augmented reality simulation session. Figure 5 shows a representative snap shot from the augmented reality based training animation.

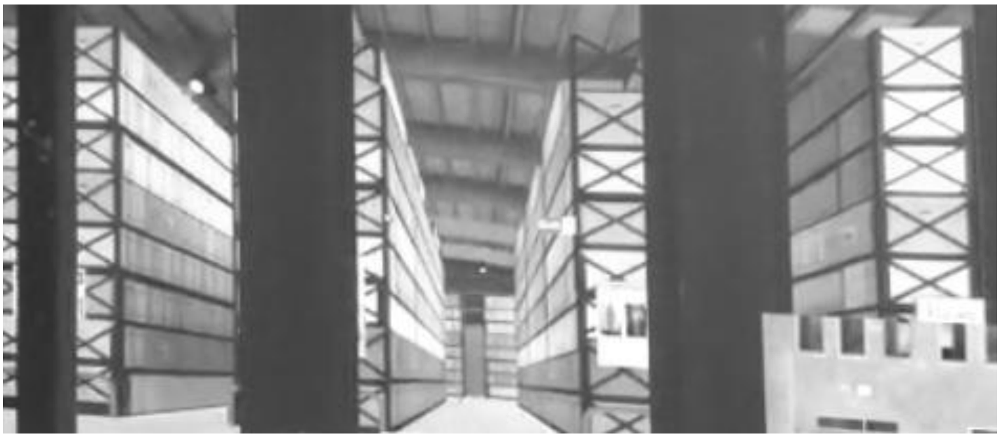


Fig. 1. Images of the industry session of VR animated video simulation

The figure shows the image of a VR animation video simulation in a realistic industry set up.

After obtaining the answers from all the participants through the questionnaire based on the recording of the results the data is fed into Artificial Neural Network (ANN) model formulated using MATLAB version 7.10.0.499.

4. RESULTS

4.1 ANN Training

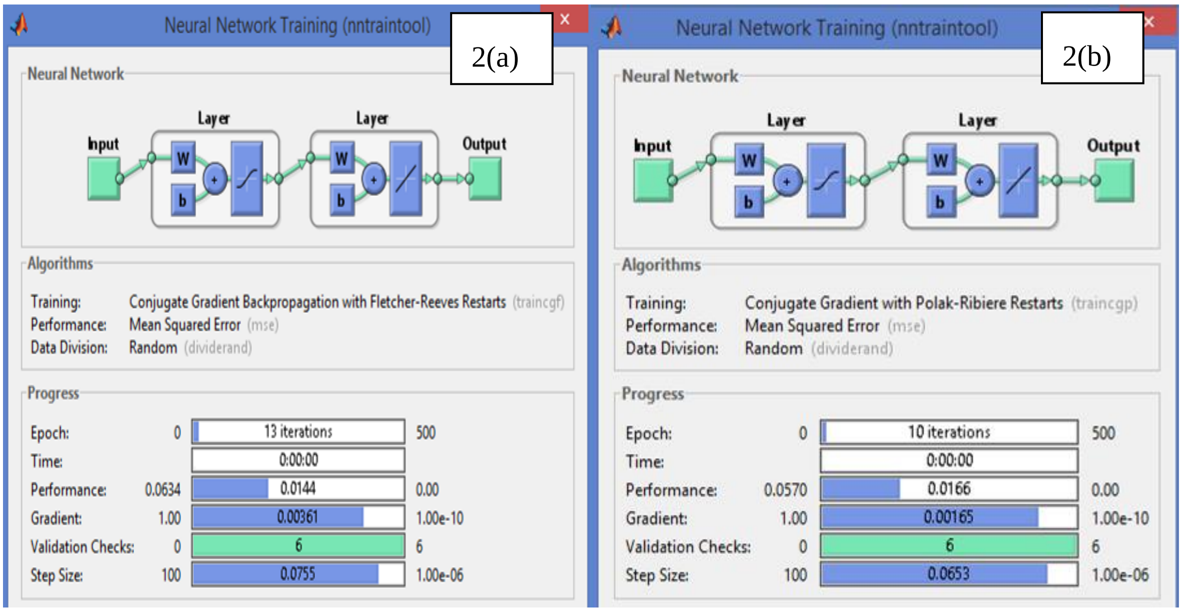
The data from the recordings of the pre-training test and the post training test is normalized for all performance indicators viz., safety, quality, productivity and cost. We have used the Backpropagation Fletcher Reeves and Polak-Ribiére updates on a comparative basis to train the ANN which to optimize the network output. While the 450 subjects pre and post training tests' recordings are being utilized to provide network training a separate set of 50 subjects are used to check the performance efficacy of the predictive outcomes of the ANN model. IN other words the test set is of the 50 pre/post subjects while the training set is of 450 subjects while the validation is subset is of 50 pre/post test values. Test set is mutually exclusive and independent of the training set of 450. LM algorithm is used for optimization of the ANN network as below.

4.2 ANN Training Results

In this section we will be describing the training behavior of the ANN for the four different key result indicators viz., safety, quality, productivity and cost using both the algorithms as described above

4.2.1 Safety Pre and post training test results normalized

Figure 2(a), (c) and (e) shows the training model, the error outcomes and the behavior of the mean square error (MSE) for the predictions carried out through the Backpropagation model with Fletcher Reeves updates. As can be seen in this training modules that the MSE minimization happens after approximately seven epochs. The R value calculated separately is found to be around 0.9 showing a good overall convergence of the outcome prediction. Figure 2 (b), (d) and (f) shows the training model, the error outcomes and the MSE for the backpropagation through the Polak Ribiere updates. In this case the MSE minimization takes place after the 4^{th} epoch. Both the models predict the safety outcomes of the experiment and the targeted MSE is around 10^{-2}.



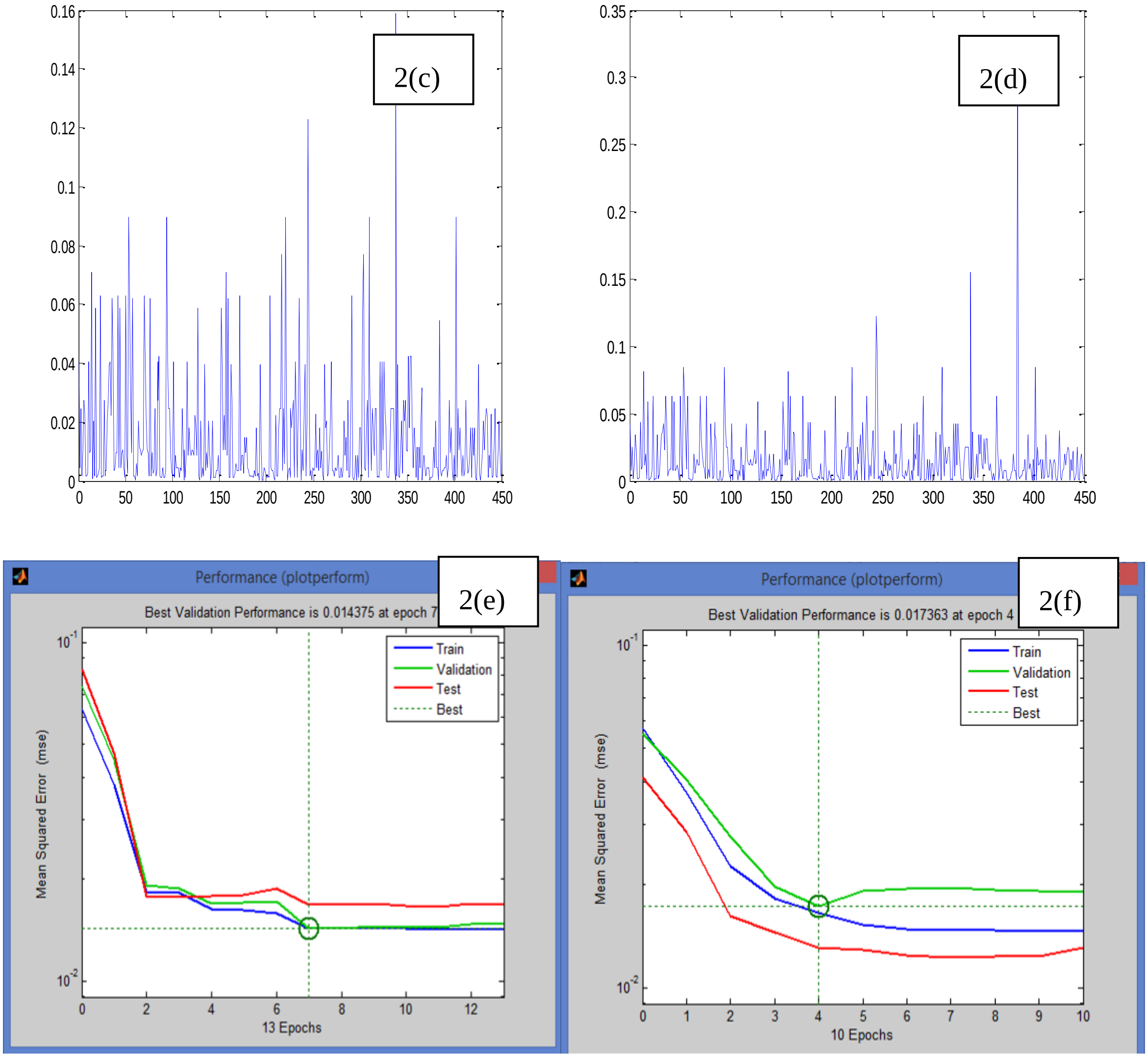


Fig. 2 (a-f). 2a, 2c, 2e for Fletcher ANN and 2b, 2d, 2f for Polak ANN Training for Safety

Using Fletcher Reeves ANN algorithm the added total error of the training, the validation and the test sets come out to be 6.6837 and for the Polak-Ribiére ANN algorithm this comes out to be 7.2168

4.2.2 Quality Pre and post training test results normalized data

Figure 3(a), (c) and (e) shows the training model, the error outcomes and the behavior of the mean square error (MSE) for the predictions carried out through the backpropagation model with Fletcher Reeves updates. As can be seen in this training modules that the MSE minimization happens after approximately four epochs. The R value calculated separately is found to be around 0.93 showing a good overall convergence of the outcome prediction. Figure 3 (b), (d) and (f) shows the training model, the error outcomes and the MSE for the backpropagation through the Polak Ribiere updates. In this case the MSE minimization takes place after the 16th epoch. Both the models predict the safety outcomes of the experiment and the targeted MSE is around 10.

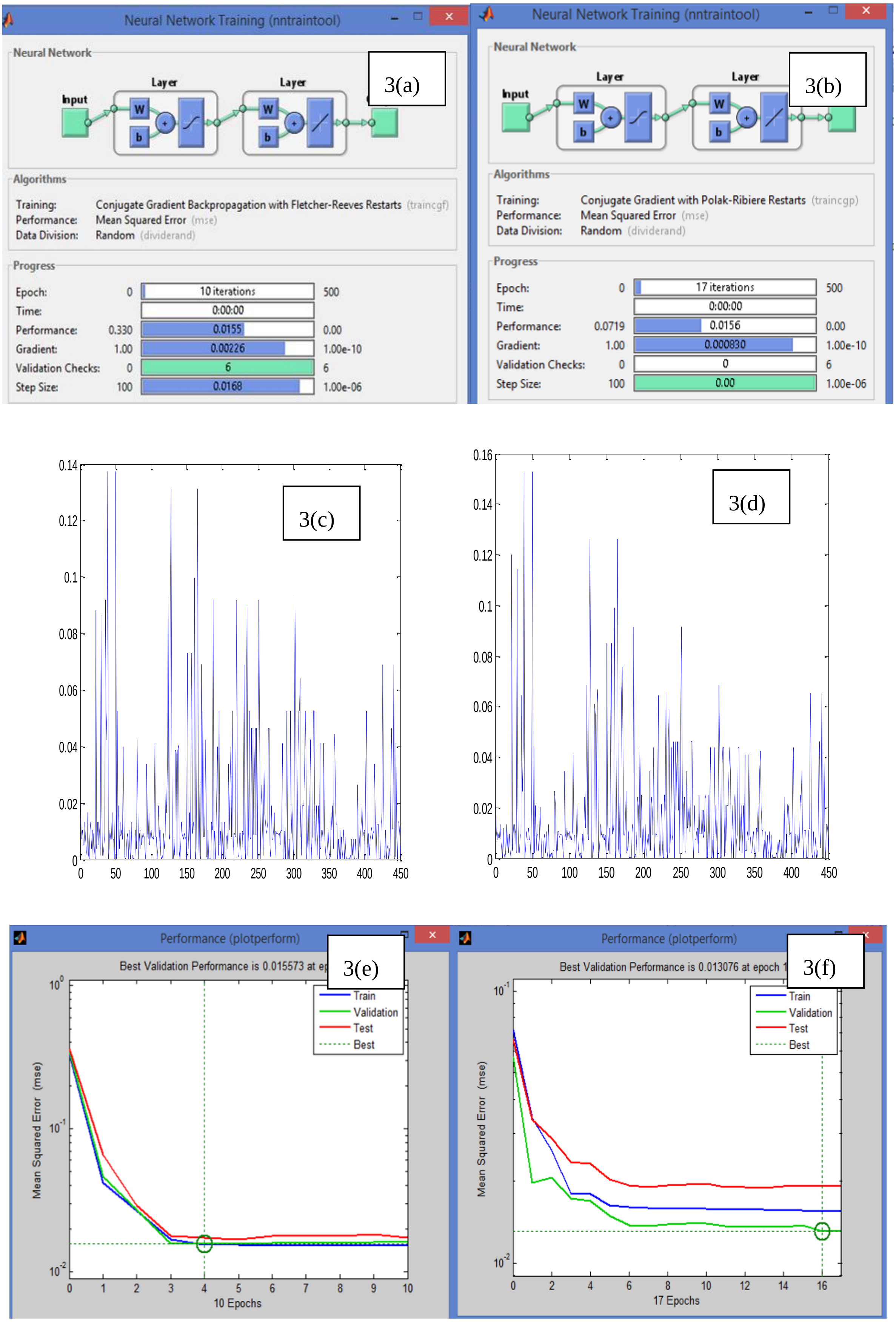
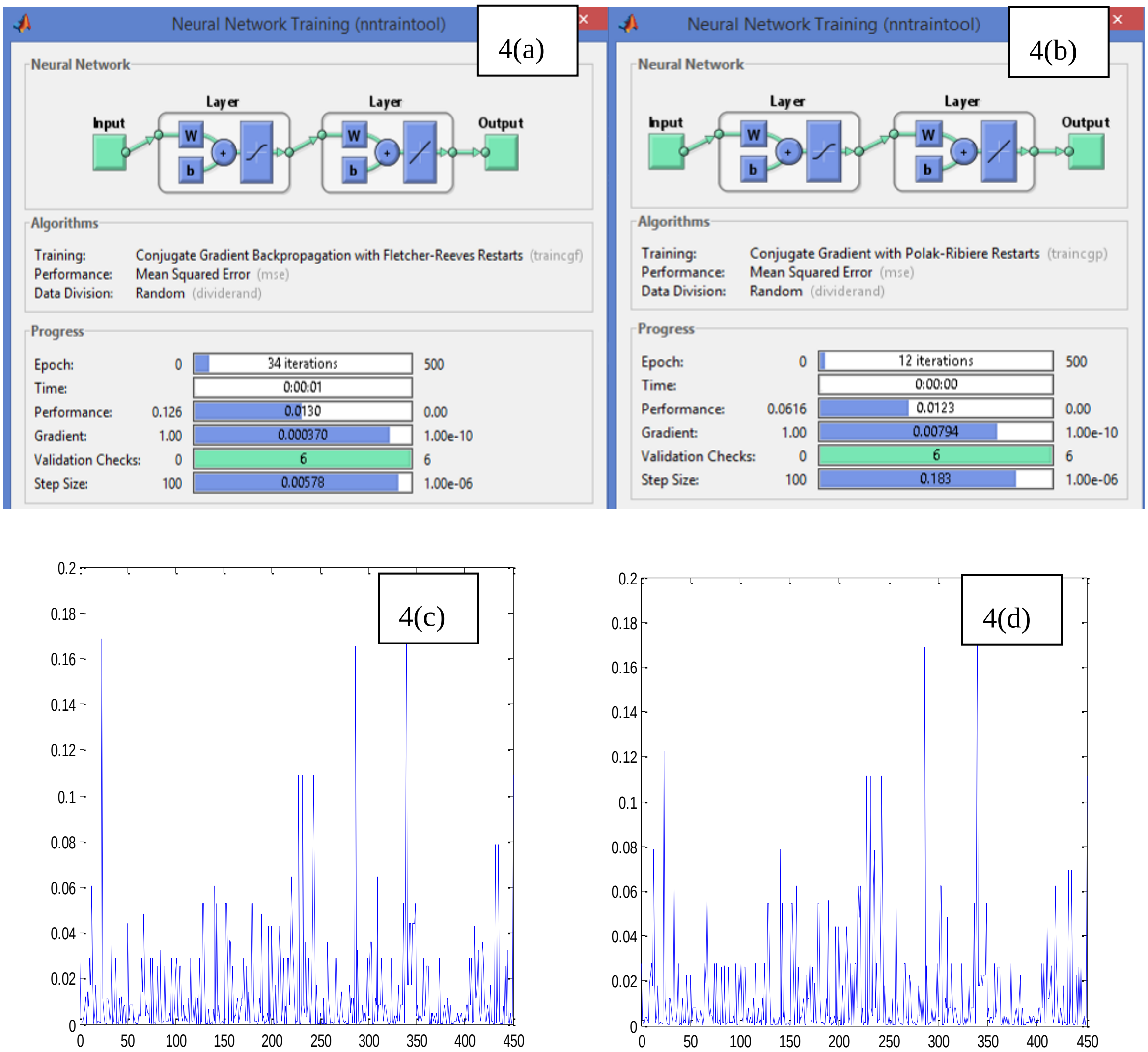


Fig. 3 (a-f). 3a, 3c, 3e for Fletcher ANN and 3b, 3d, 3f for Polak ANN Training for Quality

The added total errors of the training, the validation and the test sets using the Fletcher Reeves algorithm comes out to be 7.1391 and for that using the Polak Ribiere algorithm comes out to be 7.1157 respectively.

4.2.3 Productivity Pre and post training test results normalized

Figure 4(a), (c) and (e) shows the training model, the error outcomes and the behavior of the mean square error (MSE) for the predictions carried out through the backpropagation model with Fletcher Reeves updates. As can be seen in this training modules that the MSE minimization happens after approximately twenty eight epochs. The R value calculated separately is found to be around 0.88 showing a good overall convergence of the outcome prediction. Figure 4 (b), (d) and (f) shows the training model, the error outcomes and the MSE for the backpropagation through the Polak Ribiere updates. In this case the MSE minimization takes place after the 6th epoch. Both the models predict the safety outcomes of the experiment and the targeted MSE is around 10^{-2}.



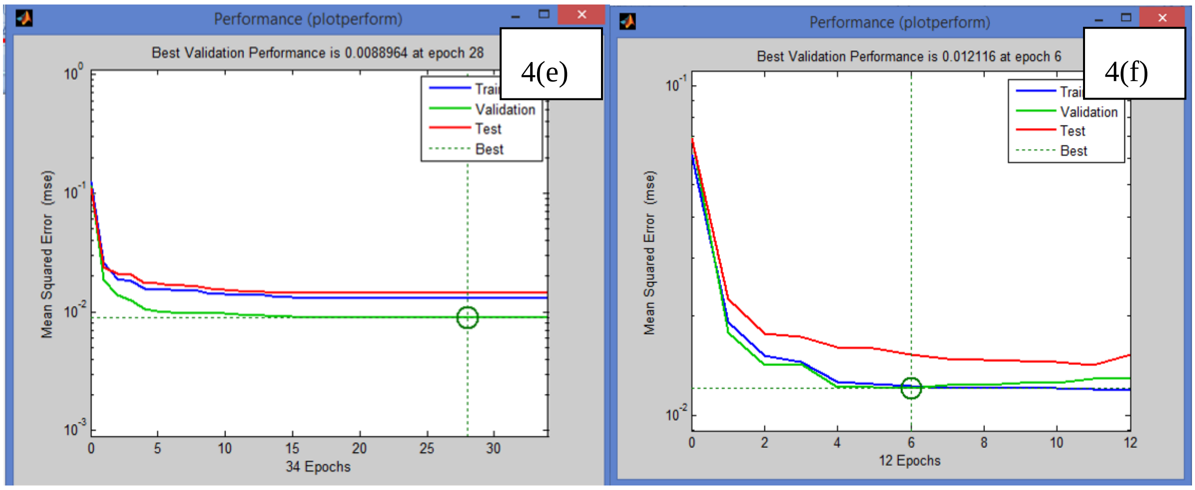
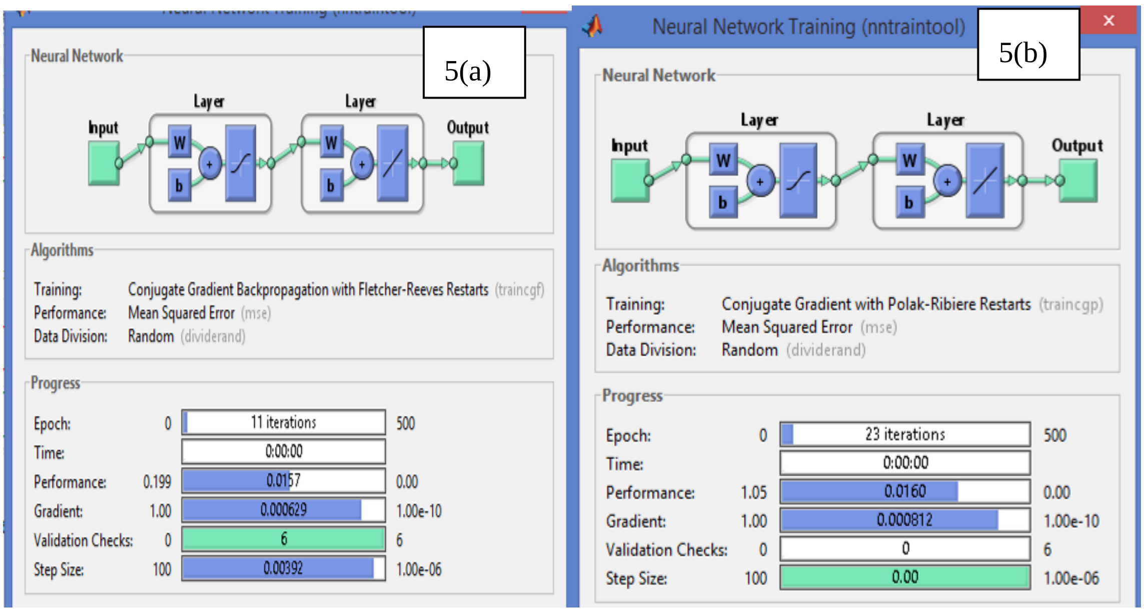


Fig. 4 (a-f). 4a,4c,4e for Fletcher ANN and 4b,4d, 4f for Polak ANN Training for Productivity

The added total errors using Fletcher Reeves algorithm for the training, the validation and the test sets come out to be 5.6129 and that using Polak Ribiere algorithm comes out to be 5.7838

4.2.4 Cost Pre and post training test results normalized

Figure 5(a), (c) and (e) shows the training model, the error outcomes and the behavior of the mean square error (MSE) for the predictions carried out through the backpropagation model with Fletcher Reeves updates. As can be seen in this training modules that the MSE minimization happens after approximately five epochs. The R value calculated separately is found to be around 0.93 showing a good overall convergence of the outcome prediction. Figure 4 (b), (d) and (f) shows the training model, the error outcomes and the MSE for the backpropagation through the Polak Ribiere updates. In this case the MSE minimization takes place after the 25^{h} epoch. Both the models predict the safety outcomes of the experiment and the targeted MSE is around 10^{-2}.



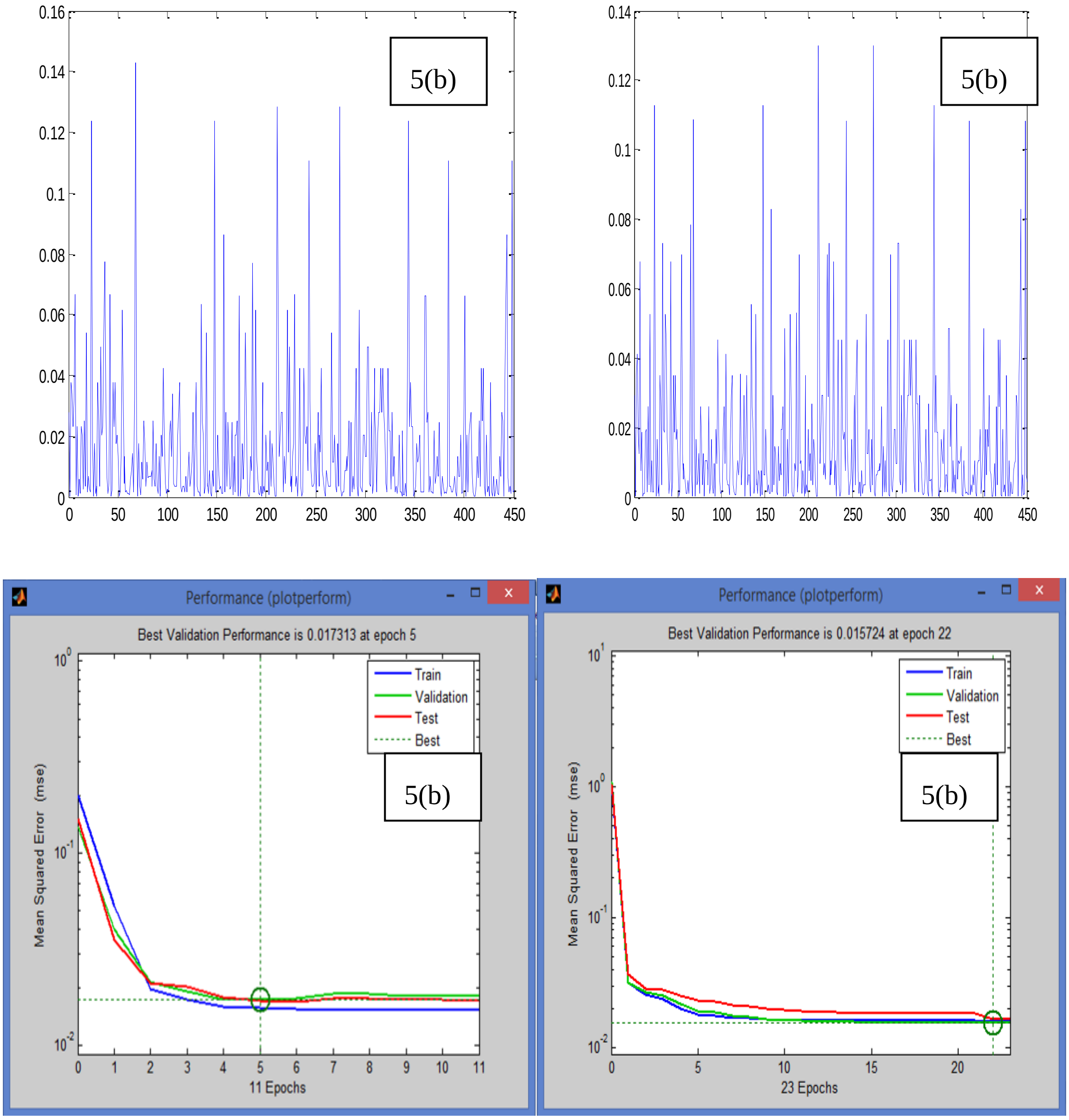


Fig. 5 (a-f). 5a,5c,5e for Fletcher ANN and 5b,5d, 5f for Polak ANN Training for Cost

The added total errors using the Fletcher Reeves ANN algorithm of the training, the validation and the test sets come out to be 7.3365 and that using the Polak-Ribiére ANN algorithm comes out to be 7.2366.

A summary of these data is made in the table 2 and 3 below which also compares the outcomes through both methods and also the processing times and no. of iterations of the different methods.

S.No	ANN Algorithm	Sum of Error-Safety	Sum of Error-Quality	Sum of Error-Productivity	Sum of Error-Cost	Best Validation at Epoch No.
1	Conjugate gradient backpropagation with Fletcher-Reeves updates	6.6837	7.1391	5.6129	7.3365	2/4/28/5 epochs respectively
	Sum error/Performance	0.0634	0.330	0.126	0.199	
2	Conjugate gradient backpropagation with Polak-Ribiére Updates	7.2168	7.1157	5.7838	7.2366	4/16/6/22 epochs respectively
	Sum error/Performance	0.0570	0.0719	0.0616	1.05	

Table 2. Sum of errors and performance comparison of Fletcher-Reeves vs. Polak-Ribiére Updates ANN

The table below shows comparison of processing time and iterations required

S.No	ANN Algorithm	Iterations – Safety	Iterations Quality	Iterations – Productivity	Iterations – Cost
1	Conjugate gradient backpropagation with Fletcher-Reeves updates	13	10	34	11
2	Conjugate gradient backpropagation with Polak-Ribiére Updates	10	17	12	23

Table 3. Comparison of Iterations and Processing time of Fletcher-Reeves vs. Polak-Ribiére Updates ANN

The conclusions of the study are that in the limited study of data coming out of a certain industry for the outcome prediction of the knowledge transfer process in warehousing settings through ANN using two different training algorithms the outcomes of the data are quite different and the Polak Ribere algorithm seems to work slightly better than the Fletcher reeves one as the no. of training epochs used for all four KPIs are slightly lower. One can say that for the data set in use the error convergence on the target is faster through use of the Polak Ribiere method. The scope of the study is limited in use of only two layers and also using unidirectional data related to a warehousing setup. Of course there network generalization for this particular outcome prediction application will need a higher order and dimension data set which is in the purview of the authors in future modules of network trainings. Also important future scope of this study is to actually visualize the different layers in terms of data convergence and what portion of the data is interpreted in what form by the individual layers which results in outcome predictions. This research will however need a lot of depth in delving into the individual raw data sets of each layer and definitely forms a direction to go ahead in future modules.

5. CONCLUSIONS

We have compared different algorithm layers with different combinations of neurons to optimize a ANN through multiple training algorithms like the Conjugate gradient Backpropagation with Fletcher reeves updates or Polak-Ribiére Updates and have mapped the efficacy of the outcome prediction ability of the network in estimating 04 KPIs associated with the warehousing virtual training module. The paper shows a new direction into how training through virtual modes can be applied to industrial problems and outcome prediction be mapped without any post training evaluation of subjects.

ACKNOWLEDGMENTS

The authors duly acknowledge IIM Lucknow and IIT Kanpur for resources related to the software packages and other support.

REFERENCES

1. Ikujiro Nonaka, Ryoko Toyoma, 2003, The Knowledge-creating theory revisted: knowledge creation as synthesizing process", Knowledge Management Research and Practice (2003) 1, pp 2-10.
2. Nonaka I (1991), "The Knowledge –Creating company" Harvard Business School Publishing , pp 96-104
3. Nonaka I and Takeuchi H (1995) "The Knowledge-creating company" Oxford university Press, New York , pp 68-94
4. Asif, M. and Searcy, C. (2014), "A composite index for measuring performance in higher education institutions", International Journal of Quality & Reliability Management, Vol. 31, No. 9, pp. 983-1001
5. Katherine M.C, E Wynn, Mei Lu Mary B.W.M (2005), "How virtual are we? Measuring virtuality and understanding its impact in a global organization", Information Systems Journal , Vol 15, Issue 4, pp 279–306
6. Maged N. ,Kamel Boulos, Lee H, Wheeler S (2007) "Second Life: an overview of the potential of 3-D virtual worlds in medical and health education", Health Information and Libraries Journal , Vol 24, Issue 4 , pp.233–245
7. Sabine W, Bockholt U, Engelke T, Gavish N , Olbrich M, Preusche C (2013), " An augmented reality training platform for assembly and maintenance skills", Robotics and Autonomous Systems, Volume 61, Issue 4, Pages 398-403
8. Al Rivas, A González-Briones, G Hernández, Javier Prieto, P Chamoso, (2020) "Artificial neural network analysis of the academic performance of students in virtual learning environments", Neurocomputing, pp1-8, https://doi.org/10.1016/j.neucom.2020.02.125
9. M.C. Alfred, D.M. Neyens, Anand K.G (2019) "Learning in simulated environments: An assessment of 4-week retention outcomes", Applied Ergonomics, Volume 74, 2019, pp 107-117, ISSN 0003-6870, https://doi.org/10.1016/j.apergo.2018.08.002.
10. R. Fletcher, C. M. Reeves, (1964) "Function minimization by conjugate gradients", The Computer Journal, Volume 7, Issue 2, , Pages 149–154
11. E. Polak, G. Ribiere, (1969), Revue française d'informatique et de recherche opérationnelle "Note sur la convergence de méthodes de directions conjuguées", Série rouge, tome 3, no R1,pp. 35-43
12. Scales, L.E., (1985) "Introduction to Non-Linear Optimization", New York, Springer-Verlag,
13. CW Choo, 1996, "The Knowing Organization: How Organizations Use Information to Construct Meaning, Create Knowledge and Make Decisions" International Journal of Information Management, Vol. 16, No. 5, pp. 329-340, 1996
14. Al-Stouhi S., Reddy C.K. (2011) Adaptive Boosting for Transfer Learning Using Dynamic Updates. In: Gunopulos D., Hofmann T., Malerba D., Vazirgiannis M. (eds) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2011. Lecture Notes in Computer Science, vol 6911. Springer, Berlin, Heidelberg
15. Shailendra Singh, Shantanu Bhattacharya, Venkataramanaiah Saddikuti (2020) "An application of AHP and fuzzy AHP with sensitivity analysis for selecting the right process to impart knowledge", International Journal of Knowledge Management Studies, Accepted, Under Publication
16. Amy J. C. Trappey, Charles V. Trappey, Tzu-An Chiang and Yi-Hsuan Huang, 2013, "Ontology-based neural network for patent knowledge management in design collaboration", International Journal of Production Research, Vol. 51, No. 7, pp 1992–2005

17. A. Singh, M. Badoni and B. Singh, "Application of least means square algorithm to shunt compensator: An experimental investigation," *2014 IEEE International Conference on Power Electronics, Drives and Energy Systems (PEDES)*, Mumbai, 2014, pp. 1-6
18. Hamid R. Nemati , David M. Steiger , Lakshmi S. Iyer , Richard T. Herschel (2002), "Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing" Decision Support Systems , Vol 33 , pp 143–161
19. Nonaka, I., Toyama, R. and Konno, N. (2000) 'SECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation", *Long Range Planning*, 33(1), pp. 1-31
20. Joshi H, Chawla D (2019) "How knowledge management influences performance? Evidences from Indian manufacturing and services firms" International Journal of Knowledge Management, Vol. 15, Issue 4, pp 56-77.

Appendix A

Algorithm for Fletcher Reeves -Conjugate gradient Backpropagation

```
%%%%one to one
 a1=k(:,9);
a2=k(:,10)'';a3=k(:,11);a4=k(:,12);
%P=[a1'; a2'; a3'; a4'];
P=[a1'];%; a2'; a3'; a4'];
 a5=k(:,13);
%take 9 and 13 together i.e. ANN training for safety post normalisation;
%take 10 and 14 together i.e. ANN training for Quality post normalisation;
%take 11 and 15 together i.e. ANN training for Productivity post normalisation;
%take 12 and 16 together i.e. ANN training for Operations post normalisation;
T=[a5'];
net = newff(P,T,5,{},'traincgf');  %%%% Fletcher Reeves -Conjugate gradient Backpropagation
 %The network is simulated and its output plotted against the targets.
Y = sim(net,P);
%plot(P,T,P,Y,'o')
%The network is trained for 500 epochs. Again the network's output is plotted.
net.trainParam.epochs = 500;
net = train(net,P,T);
Y = sim(net,P);
%plot(P,T,P,Y,'o')
err=(T-Y).^2
plot(err)
```

Appendix B

Algorithm for Polak-Ribiére updates -Conjugate gradient Backpropagation

```
%%%%one to one
a1=k(:,9);
a2=k(:,10)'';a3=k(:,11);a4=k(:,12);
%P=[a1'; a2'; a3'; a4'];
P=[a1'];%; a2'; a3'; a4'];
a5=k(:,13);
%take 9 and 13 together i.e. ANN training for safety post normalisation;
%take 10 and 14 together i.e. ANN training for Quality post normalisation;
%take 11 and 15 together i.e. ANN training for Productivity post normalisation;
%take 12 and 16 together i.e. ANN training for Operations post normalisation;
T=[a5'];
net = newff(P,T,5,{},'traincgp');  %%%% Polak-Ribiére -Conjugate gradient Backpropagation
%The network is simulated and its output plotted against the targets.
Y = sim(net,P);
%plot(P,T,P,Y,'o')
%The network is trained for 500 epochs. Again the network's output is plotted.
net.trainParam.epochs = 500;
net = train(net,P,T);
Y = sim(net,P);
%plot(P,T,P,Y,'o')
err=(T-Y).^2
plot(err)
```