Extending ANN for Optical Elements - EDFA Characteristics

V. S. Lavanya, V. K. Vaidyan

Abstract: ***Artificial Neural Network has proved to be one of the best and widely used soft computation techniques in diversified fields such as Biology, Medicine, Energy, Bioinformatics etc. Modelling in Communication has come far way forward when the industry realized its benefits over conventional method of research and development. It mainly helps in two ways. The first advantage is such that the fabrication cost or wastage is highly reduced, second being the time to final solution implementation. There are various computational methods available in market, which were effectively used in the modelling of different application in diversified fields. In this work, we will discuss how effectively we can use ANN for optical elements and extend it to address the rapid explosion of information traffic and emerging applications in communication. We consider here a basic set up of forward pumped EDFA in a WDM long haul communication system and analyze the characteristics of it through proper signaling. The characterization of the gain, and amplifier noise is again modelled with the help of ANN by appropriately using the experimental data for both modelling and testing. The simulated output from the model agrees well with the experimental data and this approach can be extended to serve as a prediction tool for designing the complex systems in optical communication. The computational time(~ms) taken to model the system and mean-square error(10-5) limited is very promising to adapt the model for future activities as desired in further modelling or fabrication of the amplifier with preferred throughput. The results of modeling envisage how favorable ANN is on building the prediction formula in optical communication networks.***

Index terms-ANN, EDFA, Modelling, Optical Amplifier

I. INTRODUCTION

Optically amplified WDM systems are economic and their reliability is both experimentally and practically proved over the recent years of outburst in optical communication [1-4]. In this context, EDFA is an ideal choice of optical amplifier where it will act as a transparent box which is insensitive to the light parameters and modulation parameters of the optical signal passing through it and also provides the simultaneous amplification for all the WDM channels [5-6]. Basic characterization of EDFA is always have scope because of the high dependency of the performance of the EDFA on the intrinsic parameters like core-radius, doping ion Er3+ concentration and their response to the source wavelength, power level, pump wavelength and pump power[7]. In order to always obtain a spectrum of interest, we must have a balanced input. It is always desirable to have the sophisticated modelling which can predict the desired values.

Revised Version Manuscript Received on May 16, 2016.

V. S. Lavanya, Department of Physics, Mar Ivanios College, Thiruvananthapuram (Kerala). India.

V. K. Vaidyan, Department of Physics, Mar Ivanios College, Thiruvananthapuram (Kerala). India.

In this work, we are trying to model a simple WDM communication system using EDFA with the help of ANN.

The uniqueness of soft computing is employed here, in learning from experimentally achieved data, generalizing in a high dimension to interpolate for the unseen inputs. There are a number of soft computing techniques available and most of them are used in one way or the other in many of the applications like handwriting recognition, automotive systems and manufacturing, image processing and data compression, architecture, decision-support systems, power systems, neuro-fuzzy systems and fuzzy logic control etc.[8-13] Soft computing is likely to witness an explosive growth in the areas of science, engineering and technology and its influence might be unpredictable.

There are quite a few techniques available in soft computing, which are already used either as a combination or used independently. Some of them are neural networks, support vector machines, genetic algorithms and fuzzy logic [14-18]. In this work, we are constrained ourselves to use neural network because of certain advantages it possesses over other techniques in the current context. The main advantage of ANN over other statistical methods is that most of them appear to assume linear relationship and normal distribution when the problem is non-linear and non-normal, whereas, ANN can model dependencies between several variables and capable of conforming to the real world[19-25]. In this work, we are trying to bring the best features of ANN to fit in the EDFA modelling. ANN are known for recognizing patterns, and, generalizing from a set of training data by making simple rules for complex problems.

II. NN IMPLEMENTATION

Neural network is the network of simple neurons. The general mathematical definition of a single neuron is as below.

$$y(x) = g\left(\sum_{i=0}^{n} w_i x_i\right) \qquad \text{------------- (1)}$$

where x is a neuron with n input dendrites (x0, x1,x2,…,xn) and one output axon y(x) and where w0,…,wn are input weights. ‘g’ is an activation function which determines based on the sum of the input. If it’s a simple threshold function, it can assume a value of either ‘0’ or ‘1’, but, always cannot hold good because of the way artificial neurons are implemented. For various reasons, it is smarter to have a smooth activation function and commonly used are threshold, sigmoid and hyperbolic tangent, and, are expressed as below,

$$g(x) = \begin{cases} 1 & if \quad x+t > 0 \\ 0 & if \quad x+t \leq 0 \end{cases} \qquad \text{------------- (2)}$$

$$g(x) = \frac{1}{1+e^{-2s(x+t)}} \qquad \text{------------- (3)}$$

$$g(x) = \tanh\big(s(x+t)\big) = \frac{\sinh(s(x+t))}{\cosh(s(x+t))}$$

$$= \frac{e^{s(x+t)} - e^{-s(x+t)}}{e^{s(x+t)} + e^{-s(x+t)}} = \frac{e^{2(s(x+t))} - 1}{e^{2(s(x+t))} + 1} \qquad \text{------------ (4)}$$

Where 't' is the value that pushes the center of the activation function away from zero and's' is a steepness parameter. In artificial neuron, it can be visualized as the incoming pulse required to activate a real neuron. During learning phase, 't' together with the weights are adjusted. Hyperbolic tangent and sigmoid activation functions are smooth with very similar graphs, the major difference in the output with hyperbolic tangent ranges from -1 to 1 and the sigmoid from 0 to1 [26].

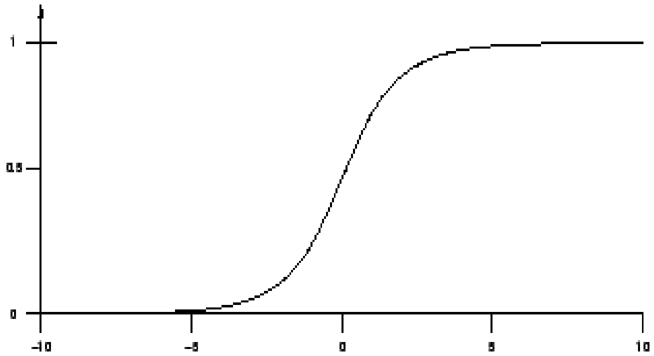

Fig 1: Sigmoid function with = 0.5 and t= 0.

An input, presented to the input layer propagated through all the layers until it reaches the output layer, where the output is returned. In a feedforward Ann, an input can be easily propagated through the network and evaluated to an output. It is more difficult to compute a clear output from a network where connections are allowed in all directions, like in the brain, since, this will create loops. Even if, recurrent networks available to code these loops of time dependencies, feedforward networks are usually a preferred choice for situations that are not time dependent [27-29]. While training the ANN with a set input and output data, the weights should be adjusted in such a way that, ANN gives the same output as given in training data. At the same time, over-fitting must be avoided such that it gives accurate results only for training data. The training process has to go through an optimization cycle so that the mean-square error is minimal. Here, Back-propagation algorithm is used, which will calculate the error after the input is propagated in the network, and the error is propagated back to the network till the weights are optimized to make the error smaller. It works in the following way,

The error e_k on a single output neuron k can be calculated as,

$$e_k = d_k - y_k$$

Where y_k is the calculated output and d_k is the desired output of neuron k.

The error value δ_k is,

$$\delta_k = e_k\, g'(y_k)$$

Where g' is the derived activation function.

The δ_j value of the proceeding layer can be calculated from the δ_k value of this layer through the below equation,

$$\delta_j = \eta\, g'(y_j) \sum_{k=0}^{K} \delta_k w_{jk}$$

Where K is the number of neurons and η is the learning rate parameter.

If we denote, Δu as the weights should be adjusted, it can be written as,

$$\Delta w_{jk} = \delta_j y_k$$

The Δuj,k value is used to adjust the weight uj,k , by uj,k =uj,k +Δuj,k and the back propagation algorithm proceed to the next input and adjusts the weights according to the output. This process goes on until a certain criteria are reached. The stop criteria are usually decided by measuring the mean square error of the training data while training with the data, and, when this mean square error reaches a desired limit, the training is stopped. This method can be extended for both training and testing data, when, more advanced stopping criteria involved [30-34]

III. EXPERIMENTAL SETUP

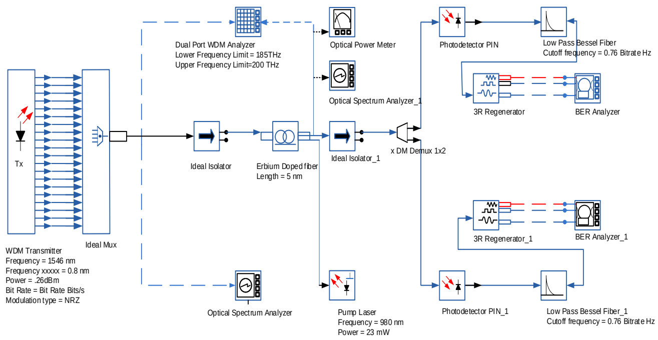

Fig 2: EDFA based optical communication system

Figure 2 shows the basic experimental set up for a WDM communication system with EDFA employed. A continuous wave laser at a launching power level of -20 dbm is used as the information source, which can be configured to emit at the wavelength from 1525 nm to 1600 nm, and is multiplexed after NRZ format modulation. The multiplexed signals are fed into a single mode fiber. The signal undergoes certain losses during its transmission in the fiber through attenuation, dispersion and non-linearity. To balance these, certain measures are introduced in the system. For example, non-linearity can be suppressed through active power management. Dispersion is compensated by using DCF fibers as shown in the setup. When the multiplexed signal are transmitted through the fiber, after a certain distance, the signal will experience attenuation, and an all optical EDFA is used to amplify the signals. Two EDFA are being used here to compensate for the attenuation losses from single mode fiber and DCF. Uni-directional passive elements like isolators used to remove unwanted signal in the backward direction. The output of such a system can be studied based on the gain performance and amplifier noise.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

IV. RESULTS AND DISCUSSION

The performance of EDFA is measured based on several characteristics results, which include Noise figure, Gain profile, ASE noise etc. In this work, we mainly focus on the modelling of certain performance measures, which makes EDFA prevalent among other optical amplifiers, and can be effectively extended for the use of other studies too. As explained earlier, the data is collected experimentally, and the ANN circuit is trained to simulate the expected results. Various studies on the performance of EDFA based on multiple conditions have been extensively studied and reported by several researchers [35-38]. Our idea here is to make a sophisticated modelling process which enables us to predict the results at a given condition, rather than waiting for the cumbersome procedures of fabrication, analysis and experiments. We have already mentioned above our choice of soft computing as ANN and how effectively we can fit this for EDFA characteristic study and extends to other components of optical communication system. As explained earlier, the data is collected experimentally, and the ANN circuit is trained to simulate the expected results. Experiments have been conducted as shown in figure to collect the necessary data for this analytic study. These data have been used to train the network and test the network, and finally the trained network will be used for predicting the behavior of unforeseen data. Figure shows the basic steps involved in the training and testing phase in ANN using neural network tool in Matlab. As we discussed earlier, a two layer feed-forward network with sigmoid hidden neurons and linear output neurons can fit multi-dimensional mapping problems arbitrarily well, given data consistency and adequate neurons in the hidden layer.

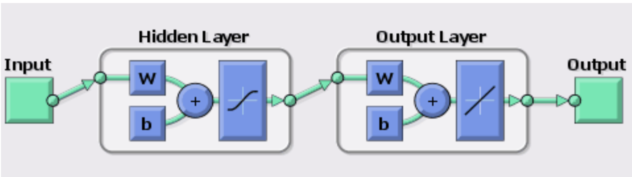

Fig 3: Two layer feed-forward NN

The number of neurons in the hidden layer can be set in the network architecture before training and can always change it in the panel if the network does not perform well after training. A training algorithm needs to be selected and different training can yield slightly different results because of the change in initial conditions and sampling. Levenberg-marquardt is a widely used algorithm which takes lesser time for generalization. trainlm is the network training function which updates weight and bias values according to Levenberg-Marquardt optimization. trainlm is often the fastest backpropagation algorithm in the toolbox, and is highly recommended as a first-choice supervised algorithm, although it does require more memory than other algorithms. The mean square error of validation samples helps in deciding on when to stop the generalization. Lower value of the Mean squared error, the average squared difference between outputs and targets, is preferred, while zero gives no error [39, 40]
There is another parameter called Regression R, which measures the correlation between outputs and targets. Close relationship will point to 1 whereas 0 shows a random relationship [41]. Figure 4 below shows the progress of training through neural network.

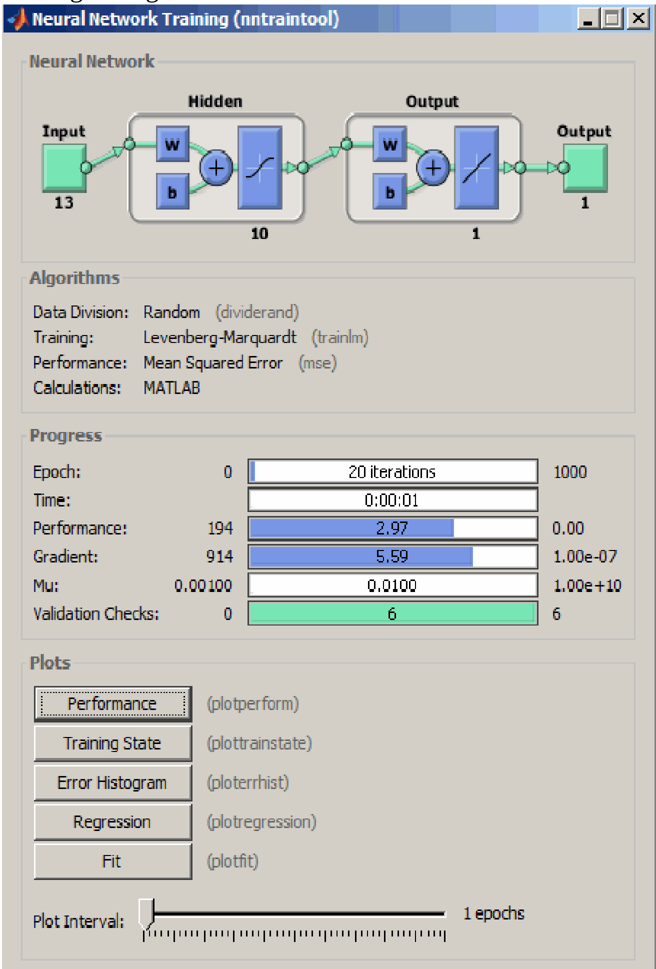

Fig 4: Neural Network Training

Figure 5 & 6 shows how the experimental characteristics curves fit in after the modelling using ANN. In the power-wavelength curve, a set of experimentally obtained data is used to model the network and the rest is used for testing the network. The result of simulation is depicted along with the experimental results, the graphs envisage how effectively the ANN modelling fitting the characteristic curves. The same method is repeated for other curves also as shown in subsequent diagram. Gain variations with wavelength and input power were of vital research topic since when the EDFA based communication system is in place. There is a significant gain tilt in within the available bandwidth of the EDFA. We do not have a flat spectra, and the gain differs at the wavelength [42-47]. Input signal power variation also play a significant role in the gain of EDFA, and how well the model fits to predict it actually helps in the designing of the WDM communication system. Similarly, noise properties are also exhibited by EDFA with respect to the input power variations and wavelength as shown in the figure.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

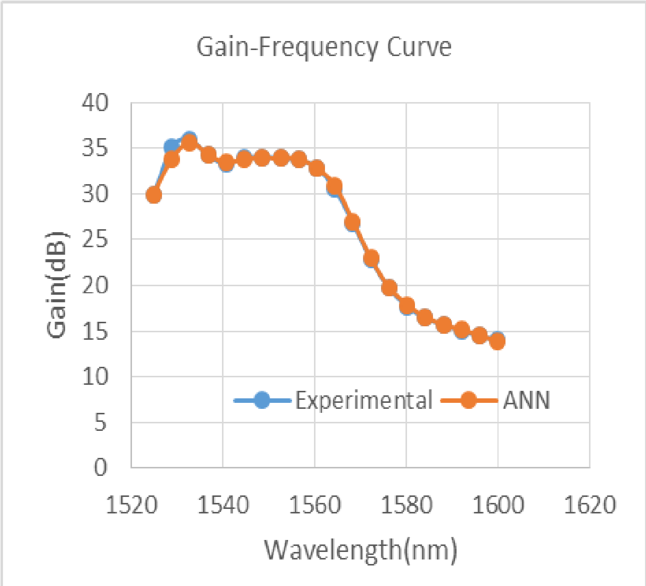

Fig 5: Gain Profile. (a)Gain-Frequency

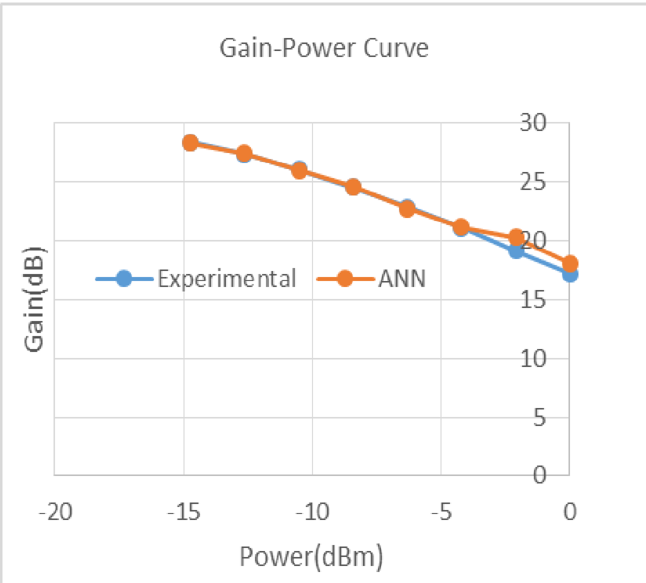

Fig 5: Gain Profile. (b) Gain-Input Power

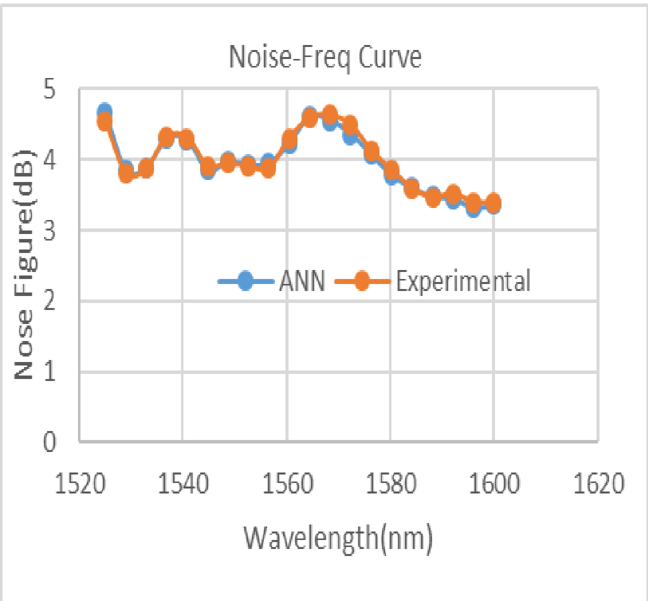

Fig 6: Noise Figure. (a)Noise-Frequency

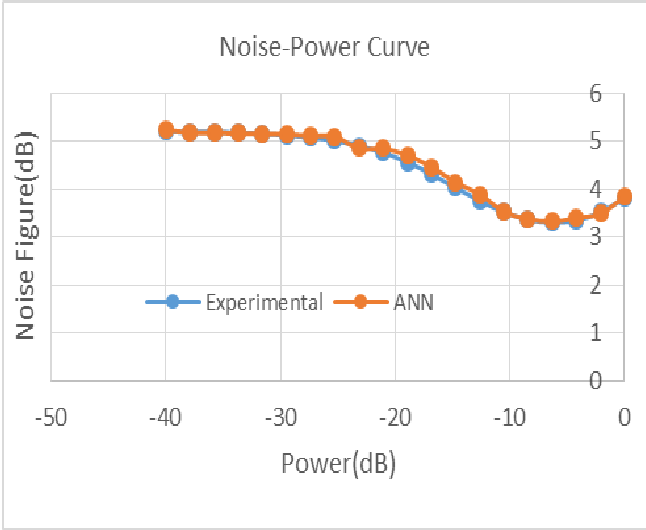

Fig 6: Noise Figure. (b) Noise-Input Power

V. CONCLUSION

In this paper, we demonstrated how ANN can be used for modelling the EDFA for the desired characteristics. For this, Gain profile and noise figure are selected as the output features of the communication network, and which can be further extended to any such measures. The theoretical results extracted from the model competes very well with the experimental data and therefore the said model can be used to design such systems. The mean square error value of the model is on the order of 10-5 with a computational time of very few milliseconds. So, EDFA characteristics is an instance of how soft computing techniques can be explored to meet the modelling requirements of optical communication networks. This is a value add for ANN due its growing demand as a successful optical amplifier. The theoretical results can be used for predicting the behavior of amplifier in any circumstances and this helps in like any other modelling tool for the early time to market and avoid overheads of test fabrications while designing the amplifier. The challenge here is to select the most suitable algorithm which can give better results and to decide the number of neurons for accurate output.

REFERENCES

1. E. Desurvire and J.R. Simpson, "Amplification of Spontaneous Emission in Erbium-Doped Single-Mode Fibers", J. Lightwave Tech., Vol.7, No.5, 835,1989.
2. P.C. Becker, N.A. Olsson, and J.R. Simpson, "Erbium-Doped Fiber Amplifiers", Academic Press, 1999.
3. G.P Agrwwal, "Fiber-Optic Communication Systems", Wiley Interscience, 3rd ed., 2002.
4. M. Melo, O. Frazao, A.L.J. Teixeira, L.A. Gomes, J.R. Ferreira D. Rocha, H.M. Salgado, "Tunable L-band erbium-doped fiber ring laser by means of induced cavity loss using a fiber taper", Applied Physics B, Vol.77, 139, 2003.
5. R. J. Mears, S. R. Baker, "Erbium Fiber Amplifiers and Lasers", Optical and Quantum Electronics, Vol.24, 517, 1992.
6. E. Desurvire, J. Simpson, and P.C. Becker, High-gain erbium-doped traveling-wave fiber amplifier," Optics Letters, vol. 12, No. 11, 1987, pp. 888–890
7. S. Lavanya, v. K. Vaidyan, "optimized flattened gain spectrum in c –band wdm using automatic gain control in bi-directionally pumped EDFA", INTERNATIONAL Journal of Engineering Research & Technology, Vol-4, No-10, Pages: 430 - 434, October 2015

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

8. J. Hertz, A. Krogh, and R.G. Palmer, Introduction to the "Theory Of Neural Computation, Addison-Wesley, Reading, Mass., 1991.
9. S. Haykin, Neural Networks: A Comprehensive Foundation, MacMillan College Publishing Co., New York, 1994.
10. W.S. McCulloch and W. Pitts, "A Logical Calculus of Ideas Immanent in Nervous Activity," Bull. Mathematical Bio-
11. R. Rosenblatt, Principles of Neurodynamics, Spartan Books, New York, 1962.
12. M. Minsky and S. Papert, Perceptrons: An lntroduction to Computational Geometry, MIT Press, Cambridge, Mass., 1969.
13. J.J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," in Roc. Nat'l Academy of Sciences, USA 79,1982, pp. 2,5542,558.
14. P. Werbos, "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences," PhD thesis, Dept. of Applied Mathematics, Harvard University, Cambridge, Mass.,1974.
15. D.E. Rumelhart and J.L. McClelland, Parallel Distributed Processing: Exploration in the Microstructure of Cognition, MIT Press, Cambridge, Mass., 1986.
16. J.A. Anderson and E. Rosenfeld, Neurocomputing: Foundations of Research, MIT Press, Cambridge, Mass., 1988.
17. S. Brunak and B. Lautrup, Neural Networks, Computers with Intuition, World Scientific, Singapore, 1990.
18. J. Eeldman, M.A. Fanty, and N.H. Goddard, "Computing with Structured Neural Networks," Computer, Vol. 21, No. 3, Mar. 1988, pp. 91-103.
19. D.O. Hebb, The OrganizationofBehavior, JohnWiley&Sons, New York, 1949.
20. R.P. Lippmann, "An Introduction to Computing with Neural Nets,"lEEEASSP Magazine, Vol. 4, No. 2, Apr. 1987, pp. 4-22.
21. A.K. Jain and J. Mao, "Neural Networks and Pattern Recognition," in Computational Intelligence: Imitating Life, J.M. Zurada, R. J. Marks 11, and C.J. Robinson, eds., IEEE Press, Piscataway, N.J., 1994, pp. 194-212.
22. T. Kohonen, Self Organization and Associative Memory, Third Edition, Springer-Verlag, New York, 1989.
23. G.A. Carpenter and S. Grossberg, Pattern Recognition by Self Organizing Neural Networks, MIT Press, Cambridge, Mass., 1991.
24. "The First Census Optical Character Recognition System Conference," R.A. Wilkinson et al., eds., . Tech. Report, NISTIR 4912, US Dept. Commerce, NIST, Gaithersburg, Md., 1992.
25. V. S. Lavanya, V. K. Vaidyan, "ANN based model of automatically gain controlled EDFA in WDM systems", J Optoelectronics and Advanced Materials, Vol-17, No-11-12, Pages: 1772 - 1777, December 2015
26. Steffen Nissen, Implementation of a fast artificial neural network library (FANN), Department of Computer Science, University of Copenhagen (DIKU) October 31, 2003
27. Martin Davis and Hillary Putnam. A computing procedure for quantification theory. ACM, 7:201–215, 1960.
28. John W. Dawson. G¨odel and the origins of computer science. In A. Beckmann, U. Berger, B. L¨owe, and J.V. Tucker, editors, Logical Approaches to Computational Barriers, CiE'06, volume 3988 of LNCS, pages 133–137, 2006.
29. F. Debart, P. Enjabert, and M. Lescot. Multimodal logic programming using equational and order-sorted logic. Theoretical computer science, 105(1):141–166, 1992.
30. Angluin, Frazier and Pitt (1990): D. Angluin, M. Frazier and L. Pitt, Learning conjunctions of Horn clauses. In Proceedings of the Thirtieth-First IEEE Symposium on Foundations of Computer Science, IEEE Computer Society Press, Washington DC.
31. Anthony and Biggs (1992): M. Anthony and N. Biggs, Computational Learning Theory: an Introduction, Cambridge University Press.
32. Bartlett (1992): P.L. Bartlett, Lower bounds on the Vapnik-Chervonenkis Dimension of multi-layer threshold networks. Technical report IML92/3, Intelligent Machines Laboratory, Department of Electrical Engineering and Computer Engineering, University of Queensland, Qld 4072, Australia, September 1992.
33. Baum (1990): E.B. Baum, Polynomial time algorithms for learning neural nets. In Proceedings of the Third Workshop on Computational Learning Theory. Morgan Kaufmann, San Mateo, CA.
34. Arthur d'Avila Garcez, Gerson Zaverucha, and Luis A.V. de Carvalho. Logical inference and inductive learning in artificial neural networks. In C. Hermann, F. Reine, and A. Strohmaier, editors, Knowledge Representation in Neural Networks, pages 33–46. Logos Verlag, Berlin, 1997.
35. Saleh, A. A. M., R. M. Jopson, J. D. Evankow, and J. Aspell, Modeling of gain in erbium-doped fiber amplifiers, IEEE Photon. Technol. Lett., Vol. 2, No. 10, 714 - 717, 1990.
36. Giles, C. R. and E. Desurvire, Modeling erbium-doped fiber amplifiers, J. Lightwave Technol., Vol. 9, No. 2, 271 - 283, 1991.
37. Lu, Y. B. and P. L. Chu, Gain flattening by using dual-core fiber in erbium-doped fiber amplifier, IEEE Photon. Technol. Lett., Vol. 12, No. 12, 1616 - 1617, 2000.
38. Martin, J. C., Erbium transversal distribution influence on the effectiveness of a doped fiber: Optimization of its performance, Opt. Commun., Vol. 194, 331 - 339, 2001.
39. R. Beale and T. Jackson ,Neural Computing - an introduction, Physics Publishing 1990
40. J. David Bolter, Turing's Man - Western culture in the computer age, Duckworth 1984
41. Alison Cawsey, Artificial Intelligence - The essence of, Prentice Hall 1998
42. Cheng, C. and M. Xiao, Optimization of an erbium-doped fiber amplifier with radial effects, Opt. Commun., Vol. 254, 215 - 222, 2005.
43. Cheng, C. and M. Xiao, Optimization of a dual pumped L-band erbium-doped fiber amplifier by genetic algorithm, J. Lightwave Technol., Vol. 24, No. 10, 3824 - 3829, 2006.
44. Chang, C. L., L. Wang, and Y. J. Chiang, A dual pumped double- pass L-band EDFA with high gain and low noise, Opt. Commun., Vol. 267, 108 - 112, 2006.
45. Choi, B. H., H. H. Park, and M. J. Chu, New pumped wavelength of 1540-nm band for long-wavelength-band erbium-doped fiber amplifier (L-band EDFA), J. Quantum Electron., Vol. 39, No. 10, 1272 - 1280, 2003.
46. Yeh, C. H., C. C. Lee, and S. Chi, S- plus C-band erbium-doped fiber amplifier in parallel structure, Opt. Commun., Vol. 241, 443 - 447, 2004.
47. Singh, R., Sunanda, and E. K. Sharma, Gain flattening by long period gratings in erbium doped fibers, Opt. Commun., Vol. 240, 123 - 132, 2004.

V S Lavanya, received MSc degree in Physics and M Tech degree in Optoelectronics from the University of Kerala in 2003 and 2006 respectively. She is presently a PhD student at the University of Kerala. Her current research interests include high speed fiber optic links for different platforms, Performance optimization of optical amplifiers, wireless communication and Internet of Things. She is the author of a number of publications in modelling amplifiers with optimized performance.

V. K. Vaidyan received his PhD from IIT Madras. He was head of Dept. of Physics in the University of Kerala. His research interests include materials science, crystal growth, thin films, composite materials, non-destructive evaluation of materials etc. His significant contributions include a new theory for Liesegang phenomena, a new method for the growth of coinage metal crystals, growth and characterization of biological crystals, to name a few. Dr. Vaidyan's research contributions are more than 100 publications and was a recipient of Commonwealth Academic Staff Fellowship in 1983 and ABI Gold Record of Achievement in 1997. Dr. Vaidyan has served on the Academic bodies of several Indian Universities and has been a member of various scientific and professional organizations. He was the Visiting Fellow of the University of Manchester Institute of Science & Technology (UMIST), UK during 1983-84 and is a of Fellow and Chartered Physicist of Institute of Physics (London).

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.