Volume 3, No. 3, May-June 2012

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Nonlinear Regression of characteristic of gunn diode: A Neurocomputing Approach

Mrs. S. N. Kale
Asstt. Professor, Dept. of Applied Electronics,
Sant Gadge Baba Amravati University,
Amravati, India
sujatankale@rediffmail.com

Dr. S.V.Dudul
Professor & Head, Dept. of Applied Electronics,
Sant Gadge Baba Amravati University,
Amravati, India
dudulsv@rediffmail.com

Abstract: Non-linearity is observed in the transfer characteristics of gunn diode. Neural Network can elegantly solve a typical regression of characteristic of gunn diode. The dataset is obtained by performing experiment on a typical continuous wave gunn diode MA49156-30. The numbers of readings are regarded as samples. The dataset is obtained, which is used for regression. After rigorous computer simulations authors develop an optimal MLP NN model, which elegantly performs such a nonlinear regression. Results show that the proposed optimal MLP NN model has a MSE as low as 4.24 x 10^{-5}, correlation coefficient as high as 0.9852 when it is validated on the cross validation dataset.

Keywords: Regression, MLP NN, RBF NN, Jordan Elman network, Statistical model.

I. INTRODUCTION

Literature survey [1, 2, 3, 4] shows that Neural Networks (NN) have been effectively used for nonlinear multivariable regression. However, there is still enough scope to choose an appropriate NN model so that the performance measures are optimized to approach zero and unity for MSE (mean square error) and correlation coefficient (r), respectively. In regression, both the input data and desired response are experimental variables (normally real numbers) created by a single unknown underlying mechanism. The goal in regression is to find the parameters of the best linear approximation to the input and the desired response pairs. In multivariable nonlinear regression, conventional techniques such as least square approach generally do not work reasonably [5]. Therefore NN approach is worth considering for solving nonlinear multivariable regression problem [6]. In a gunn diode V-I (voltage-current) characteristic non-linearity is observed in the characteristic. The data is obtained by performing experiment on a typical continuous wave gunn diode MA49156-30. The number of readings is treated as samples.

Optimal MLP NN (Multilayer Perceptron Neural Network) is developed for regression of gunn diode characteristic. Other NN configuration such as RBF (Radial Basis Function) and Jordan Elman Neural Network have also been considered for this regression.

Statistical models are also developed using conventional technique.

This paper deals with the multivariable regression using NN approach. Here a dataset is obtained, which is used for regression. As shown in Table 1, there are 96 training patterns out of which (70%) samples are used to train the NN model and (30%) different independent samples are used to assess the performance of an estimated network model as shown in Table I.

Table 1- Gunn diode dataset used for NN based Model

No.of total samples	No. of training Samples(70%)	No. of cross validation Samples (30%)
96	67	29

Independent validation method in statistics is used to evaluate the NN in which the available data are divided into a training set and a test set. The training data is used to update the weights, in the network. The test data are then used to assess how well the network has generalized. The learning and generalization ability of the estimated NN model is assessed on the basis of performance measures such as MSE, NMSE (normalized mean square error) and correlation coefficient, r.
Fig.1 exhibits the transfer characteristics of typical electronic device, gunn diode.

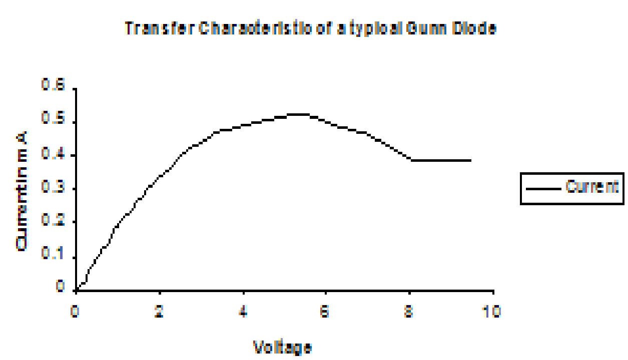

Figure.1 Transfer Characteristic of a Typical Gunn Diode MA49156-30

II. COMPUTER SIMULATION

A. MLP NN:

MLP based NN model is used in this study because it has solid theoretical foundation. The main reason for this is its ability to model simple as well as very complex functional relationship. This has been proven through a large number of

CONFERENCE PAPER
National Conference on Information and Communication Technology for Development
Organized by PRMITR, Amravati (MS) India
http://mitra.ac.in/forthcoming.html

practical applications [7]. It is shown that all continuous function can be approximated to any desired accuracy in terms of the uniform norm with a network of one hidden layer of sigmoidal or hyperbolic tangent, hidden units as well as output unit [8]. MLPs are feedforward Neural Networks trained with the standard backpropagation algorithm. They are supervised networks so they require a desired response to be trained. The configuration of MLP NN is determined by number of hidden layers, number of the neurons in each of hidden layer as well as the type of activation function used for the neurons. Backpropagation algorithm with momentum is an improvement to the straight gradient-descent search in the sense that a memory term (the past increment to the weight) is used to speed up and stabilize convergence.

Following Table II shows various parameters of the MLP NN model which are varied for obtaining optimal parameters.

Supervised learning epochs= 1000, Error threshold = 0.01, Transfer function in hidden layer= tanh, No. of PEs in input layer = 01, No. of PEs in output layer =1, No. of connection weights for 1-2-1 architecture (P) =7, Total no. of exemplars in training dataset (N) = 67, N/P = 9.5

Table 2 – Variable parameters of MLP NN Model

S.N.	Parameter	Typical Range	Optimal parameter
1	Hidden Layer	1 to 4	**1**
2	PE	1 to 20	**2**
3	Learning Rule	Momentum (Mom), Conjugate gradient (CG), Levenberg Marquardt (LM), Quick propagation (QP), Step, Delta bar delta	**Levenberg Marquardt (LM)**
4	Transfer Function in output layer	Linear, Lineartanh, Tanh	**Linear**

Here number of hidden layer is varied from 1 to 4 and performance measures of the MLP NN model are found better for single hidden layer as shown in Table III. With increase in number of hidden layers the performance of the network has not improved significantly.

Table 3– Number of Hidden layer and r (correlation coefficient)

S.N.	No. of Hidden layer	r
1	**1**	**0.9852**
2	2	0.9800
3	3	0.9642
4	4	0.9628

PEs are varied from 1 to 20 for the hidden layer. Optimal values of MSE (Mean Square Error) and r-correlation coefficient are obtained for 2 PEs in the single hidden layer.

Figure 2 gives regression capability of MLP NN on cross validation dataset**,** which portrays desired output and actual output of the MLP NN on cross validation data set. It is seen that actual output follows the desired output very closely.

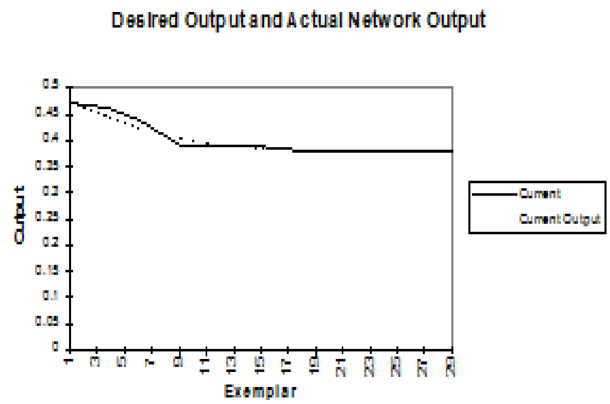

Figure. 2 Regression capability of MLP NN on cross validation Dataset

For the given dataset MLP NN model is trained for five times. The performance measures such as MSE, NMSE (Normalised Mean Square Error) and r on training dataset and testing dataset are obtained. The correlation coefficient on test dataset is found as high as 0.9852, MSE = 4.24 x 10^{-5} and NMSE =0.0434

B. RBF NN (Radial Basis Function):

A design of NN can be viewed as a curve-fitting (approximation) problem in a high-dimensional space. According to this viewpoint, learning is equivalent to finding a surface in a multidimensional space that provides a best fit to the training data, with the criterion for "best fit" being measured in some statistical sense. In the context of a NN, the hidden units provide a set of "functions" that constitute an arbitrary "basis" for the input patterns (vectors) when they are expanded into the hidden space; these functions are called RBF. RBF were first introduced in the solution of the real multivariate interpolation problem [9, 10].

A mathematical justification for the rationale of a nonlinear transformation followed by a linear transformation may be traced back to an early paper by Cover [11]. Another important point is the fact that the dimension of the hidden space is directly related to the capacity of the network to approximate a smooth input-output mapping [12,13]; the higher the dimension of the hidden space, the more accurate the approximation will be. The variable parameters of RBF NN are listed in Table IV.

Table 4 – Variable parameters of RBF NN Model

S.N.	Parameter	Typical Range	Optimal Parameter
1	Cluster centers	05-67	05
2	Unsupervised learning rule	Conscience-full, standard full	Conscience-full
3	Supervised learning rule	Momentum, Conjugate gradient, Levenberg Marquardt, Quick Propagation, Step, delta bar delta	Levenberg Marquardt
4	Metric	Euclidean, Dot Product, Box car	Euclidean
5	Transfer function in output layer	Linear, Lineartanh, Tanh	lineartanh

CONFERENCE PAPER
National Conference on Information and Communication Technology for Development
Organized by PRMITR, Amravati (MS) India
http://mitra.ac.in/forthcoming.html

Number of cluster centers is varied from 5 to 67 and optimal performance is obtained at cluster centers as 05.

In unsupervised learning Conscience-full and standard full algorithms are used. For conscience full competitive learning MSE is obtained as low as 0.00061and r as high as 0.9567.

In Supervised learning, similar to the MLP NN computer simulation the learning rules are varied and r is optimum for levenberg marquardt learning rule. Competitive learning metric is varied as Euclidean, Dot Product and Box car and Euclidean metric performs optimally.

Transfer function in output layer is varied as Linear, Lineartanh and Tanh. It is found that best results are obtained for levenberg marquardt learning rule with 05 cluster centers and lineartanh transfer function in output layer.

For the given dataset RBF NN model is trained for five times. The performance measures like MSE, NMSE and r on training dataset and testing dataset are obtained. The correlation coefficient on test dataset is found as 0.9567, MSE = 0.00061and NMSE = 0.6349.

Figure 3 gives regression capability of RBF NN that shows desired output and actual output of the RBF NN on cross validation data set. It is seen that actual output follows the desired output.

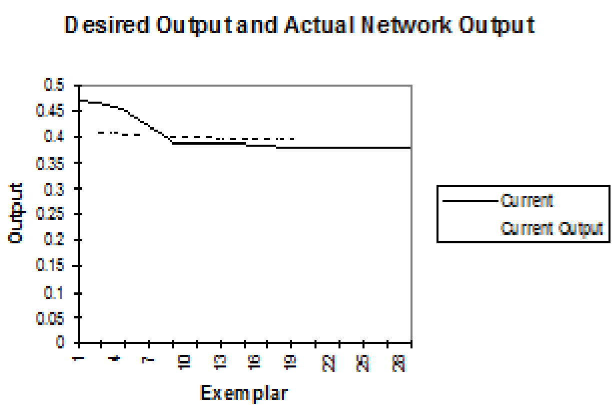

Figure. 3 Regression capability of RBF NN on cross validation Dataset

C. Jordan Elman Neural Network:

Recurrent networks are neural networks with one or more feedback loops. The Recurrent networks are used as input-output mapping networks and also as associative memories [14]. By definition, the input space of a mapping network is mapped onto an output space, a recurrent network responds temporarily to an externally applied input signal. Recurrent networks can be considered as dynamically driven recurrent networks. Because of global feedback memory requirement reduces significantly [15].

Jordan and Elman networks extend the multilayer perceptron with context units, which are processing elements (PEs) that remember past activity. Context units provide the network with the ability to extract temporal information from the data. Following Table V shows various parameters of the Jordan Elman NN model, which are varied for obtaining optimal parameters. Same architecture is used for Jordan Elman NN as that of MLP.

Table 5-Variable parameters of Jordan Elman NN Model

S.N.	Parameter	Typical Range	Optimal parameter
1	Learning Rule	Momentum (Mom), Conjugate gradient (CG), Levenberg Marquardt (LM), Quick propagation (QP), Step, Delta bar delta	Levenberg Marquardt
2	Transfer function in output layer	Linear, Lineartanh, Tanh	Linear
3	Context Unit Transfer Function	Integrator Axon, Tanh Integrator Axon, Sigmoid Integrator Axon, Context axon, Tanh Context axon, Sigmoid Context axon	Integrator Axon
4	Context Unit time constant	0.1 to 0.9	0.8

Here also different supervising learning rules are attempted**.** It is found that the best results are obtained for Levenberg Marquardt learning rule in hidden as well as output layers. Transfer functions are varied in the output layer and optimal parameters are found for linear transfer function.

In Jordan Elman NN, context unit transfer functions are varied for the optimal performance and it is found for sigmoid integrator axon. Also time constants are varied from 0.1 to 0.9 and optimal performance is obtained for time constant 0.8. For the given dataset Jordan Elman NN model is trained for five times. The performance parameters are MSE =0.00011, NMSE = 0.1226 and r =0.9839.

Fig 4 depicts regression capability of Jordan Elman NN on cross validation dataset. It is seen that actual output barely follows the desired output.

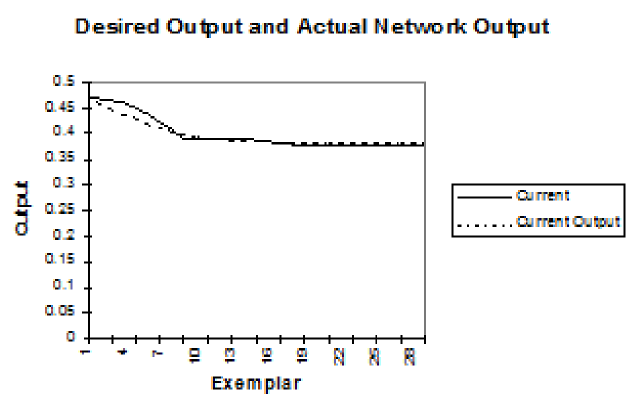

Figure. 4 Regression capability of Jordan Elman NN on Testing Dataset

The optimal performance of architectures of every NN is shown in Table VI. Levenberg Marquardt learning rule is used in hidden as well as output layers. The optimal performance is obtained for MLP NN with MSE as low as 4.24 x $10^{-5}$and r as high as 0.9852. It is also observed that for MLP NN the time required for training the network per epoch per exemplar is 14.9 microseconds. Though the percentage error is sensibly less in Jordan Elman NN the performance of MLP NN is superior as compared to other networks.

CONFERENCE PAPER
National Conference on Information and Communication Technology for Development
Organized by PRMITR, Amravati (MS) India
http://mitra.ac.in/forthcoming.html

Table 6 - Comparison of all the NN Architectures on Test ***dataset***

NN	**Transfer Function (Output layer)**	**Learning Rule**	**MSE**	**r**	**% Error**	**Time/epoch/ Exemplar in microseconds**
MLP (1-2-1)	Linear	LM	4.24×10^{-5}	0.9852	1.76	14.9
RBF	Lineartanh	LM	0.00061	0.9567	4.88	14.9
Jordan Elman (1-2-1)	Linear	LM	0.00011	0.9839	2.01	14.9

Table VII displays the regression performance of various neural networks. The performance parameters are optimum for MLP NN.

Table 7- Regression performance of NN models

S.N.	**NN**	**Performance Measures on Cross Validation Dataset**		
		MSE	**NMSE**	**r**
1	MLP	4.24×10^{-5}	0.0434	0.9852
2	RBF	0.00061	0.6349	0.9567
3	Jordan Elman	0.00011	0.1226	0.9839

D. Conventional approach:

Statistical models are developed using conventional technique. Linear, nonlinear, nonparametric and PLSR (partial leas square regression) models are built for given dataset. Table VIII shows the MSE for statistical models using various functions for Gunn diode. MSE is minimum for linear model equal to 4.53×10^{-03} on training and 0.045528 for CV dataset. In case of nonlinear regressor, minimum MSE is obtained for Function Y_1 (Degree 3) on cross validation samples. Nonparametric regressor with median method has MSE on train approaching 6.70×10^{-04} and on CV it is 3.66×10^{-03}. In PLSR model MSE on CV is 0.0105845 and on train is 0.00453. It is found that the least MSE is obtained for nonlinear model.

Table 8 - MSE for statistical models of Gunn diode

S.N.	**Model for Gunn diode dataset**			**MSE**	
				Train	**CV**
1	Linear Regression			4.53E-03	0.045528
2	Nonlinear regression	Function	Degree		
		Y_1	1	4.53E-03	0.105184
			2	1.82E-04	0.012962
			3	1.34E-04	1.13E-03
			4	1.31E-04	7.65E-03
			5	9.08E-05	1.329614
			6	8.95E-05	3.135899
			7	6.47E-05	46.4979
			8	3.90E-05	1235.07
			9	3.48E-05	0.085667
			10	2.01E-05	55263.66
		Y_2	--	unknown error occurred	
		Y_3	--	72.943	6.82E-03
		Y_4	--	7.22E-03	0.158928
		Y_5	--	3.01E-04	0.01712
		Y_6	--	3.78E-04	0.018823
		Y_7	--	5.67E+03	0.097242
		Y_8	--	3.27E-04	0.01033
		Y_9	--	can not fit function	
		Y_{10}	--	2.45E-04	0.016632
		Y_{11}	--	1.12E-04	0.026318
		Y_{12}	--	can not fit function	
		Y_{13}	--	can not fit function	
		Y_{14}	--	1.82E-04	0.010774
		Y_{15}	--	2.05E-03	0.027015
		Y_{16}	--	3.36E-04	0.018848
		Y_{17}	--	unknown error occurred	
3	PLSR	--	--	0.00453	0.0105845
4	Nonparametric Regression	Methods	Polynomial Degree		
		Lowess	1	2.00E-04	4.77E-03
			2	1.40E-04	4.77E-03
			3	4.08E-05	0.065523
			4	4.13E-05	0.065523
		Robust Lowess	1	4.42E-04	1.69E-02
			2	2.08E-04	1.69E-02
			3	5.83E-05	0.047755
			4	5.87E-05	0.047726
		Mean	--	7.62E-04	3.77E-03
		Median	--	6.70E-04	3.66E-03
		Polynomial	1	0.006677	0.090638
			2	0.010688	3.704496
			3	6.12256	9533.8
			4	7.1326	9533.8

It is very obvious from Table IX that neural network regressor having the least MSE for train as well as CV dataset amongst all the regression techniques discussed above.

Table 9 - Comparison of regression techniques on Gunn Diode dataset

S.N.	**Regression Technique**	**MSE on**		**Function / Network**
		Train	**CV**	
1	Neural Network	0.000107	$\mathbf{4.24 \times 10^{-5}}$	MLP NN
2	Linear regression	0.004527	0.045528	--
3	Nonlinear Regression	0.000134	0.001133	Y_1 (Deg3) = pr1+pr2*X1+pr3*X1^2 +pr4*X1^3
4	Nonparametric	0.00067	0.00366	Median
5	PLSR	0.00453	0.010585	--

CONFERENCE PAPER
National Conference on Information and Communication Technology for Development
Organized by PRMITR, Amravati (MS) India
http://mitra.ac.in/forthcoming.html

III. CONCLUSION

Results show that a MLP NN comprising of one hidden layer having architecture (1-2-1) is able to solve such a nonlinear regression problem with sensible accuracy. When the performance of MLP, RBF, Jordan Elman and Modular NN based regression are carefully examined for data set, MLP NN has clearly outperformed its RBF NN and Jordan Elman counterpart with respect to the performance measures such as MSE, NMSE, and r. Correlation coefficient is obtained as 0.9852 and MSE and NMSE are found to be as low as 4.24×10^{-5} and 0.0434, respectively. Moreover, the actual output of the estimated MLP NN model follows the desired output more closely than that of other NN models. It is also seen that the time per epoch per exemplar required to train the network is sensibly less for MLP NN amongst all the networks. MLP NN has achieved low percentage error equal to 1.76 % on cross validation dataset. Performance of MLP NN is definitely better when compared with conventional statistical models. MSE is least for MLP NN on training as well as CV dataset.

Proposed MLP NN model is able to accomplish the Regression of characteristic of gunn diode using Multilayer Perceptron Neural Network, which is a major contribution of this research work.

IV. REFERENCES

[1]. S. I. Wu and H. Zheng, Stock Index Forecasting using Recurrent Neural Networks. Proc. The IASTED Conference on Artificial Intelligence and Applications, Innsbruck, Austria, 2006

[2]. Richard D. de Veaux, Jennifer Schumi, Jason Schweinsberg, Lyle H. Ungar, Prediction Intervals for Neural Networks via Nonlinear Regression, Technometrics, Vol. 40, No. 4, pp. 273-282, Nov.1998

[3]. Rahul Barman, B. Prasad Kumar, P. C. Pandey Tsunami travel time prediction using neural networks, Geophysical research letters, vol. 33, L16612, 2006

[4]. Seyed Mehdi Sadat-Hashemi et. al. ,Predicting the type of pregnancy using artificial neural networks and multinomial logistic regression: a comparison study, Journal-Neural computing and applications,Volume 14, Number 3, pp 198-202, Springer London, September, 2005

[5]. Ham M. Fredric, Ivica Kostanic, Principles of neurocomputing for science & engineering (Tata McGraw Hill)

[6]. Haykin S., Neural Networks: A Comprehensive Foundation, Pearson Education.

[7]. Demuth H. and Beale M. Neural Network Toolbox: User's Guide, Version 3.0 (The MathWorks, Inc., Natick, MA,1998)

[8]. Cybenko G. Approximation by superpositions of a sigmoidal function. Mathematics of Control,Signals, and Systems, 2(4), 303-314, 1989.

[9]. Powell M. J. D.,. "Radial basis functions for multivariable interpolation: A review," IMA Conference on Algorithms for the Approximation of Functions and Data, pp.143-167RMCS, Shrivenham, England. (1985)

[10]. Light W. Ridge functions, sigmoidal functions and Neural Networks in E.W.Cheney, C.K. Chui, and L L. Schumaker, eds, Approximation Theory VII, pp. 163-206, (Boston: Academic Press 1992).

[11]. Cover T. M., "Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition," IEEE Transactions on Electronic Computers, vol. EC-14, pp. 326-334 ,1965.

[12]. Mhaskar H. N., "NNs for optimal approximation of smooth and analytic functions," Neural Computation, vol. 8, pp. 1731-1742, 1996.

[13]. Niyogi P. and F. Girosi,. "On the relationship between generalization error, hypothesis complexity, and sample complexity for radial basis functions," Neural Computation, vol. 8, pp. 819-842, 1996

[14]. Principe J.C., Euliano N Lefebvre W.C. Neural & Adaptive Systems: Fundamentals Through Simulations (John Wiley & Sons,Inc.) 2000

[15]. Quin S. Z., Su H. T., and McAvoy T. J. Comparison of Four Neural Net Learning Methods for Dynamic System Identification. IEEE Transactions on Neural Networks, vol. 3, no. 1, pp. 122-30, 1992

CONFERENCE PAPER
National Conference on Information and Communication Technology for Development
Organized by PRMITR, Amravati (MS) India
http://mitra.ac.in/forthcoming.html