International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 3 (2015) pp. 8143-8150

http://www.ripublication.com

Evolutionary Optimization of Artificial Neural Networks in Prediction of Cancer using Mixed-Integer Programming Techniques

V.Shyamala Devi and S. Shenbaga Ezhil

Asst. Prof., Dept. of Mathematics, Sathyabama University, Chennai – 119
shyamaths@gmail.com
shenbaga_ezhil@rediff.com

Abstract

A novel application to the optimization of artificial neural networks (ANNs) is presented in this paper. In this context, the weight and architecture optimization of ANNs can be formulated as a Mixed-Integer Optimization problem. A Mixed-Integer evolutionary algorithm (Mixed-Integer Hybrid Differential Evolution, (MIHDE) is used to optimize the ANN. Finally, the optimized ANN is applied to the prediction of Colon Rectum Cancer. The satisfactory results are achieved, and demonstrate that the optimized ANN by MIHDE can effectively predict the Colon Rectum Cancer.

Keywords: Evolutionary algorithms, artificial neural networks, Mixed-integer evolutionary algorithm, Mixed-Integer hybrid differential evolution, Prediction, Artificial neural networks, Optimization.

Introduction

In recent years, Evolutionary Algorithms (EAs) have been applied to the ANN's optimization. EAs are powerful search algorithms based on the mechanism of natural selection. Unlike conventional search algorithms, they simultaneously consider many points in the search space so as to increase the chance of global convergence. Many considerable efforts at obtaining optimal ANNs based on EAs have been reported in the literature and comprehensively reviewed by Yao [1]. As a result, they customarily resort to developing specific or exclusive methods to handle such a problem [2]-[7]. We should make effort to identify the exact attribute of the optimization problem of ANN's architecture and weights and then use an appropriate optimization tool to solve it.

To demonstrate the capability to predict the optimized ANN by MIHDE [8-10] is used colon rectum cancer. Cancer refers to cells that grow larger that 2mm in every 3 months and multiply uncontrollably and spread to other parts of the body. In this paper we have developed an Artificial Neural Network (ANN) models for Colon Rectum Cancer which can be used for all type of diagnosis and detection [1]. Artificial Neural Network systems are made to learn the cancer data by the use of training algorithms. Learning involves the extraction of rules or pattern from the historic data.

Evolutionary Optimization of Anns

In this paper, a Mixed-Integer encoding scheme is used to represent the optimization parameters of ANNs, including the number of hidden layers, the number of nodes in each hidden layer, the types of node transfer functions and the connection weights. Except for the connection weights, the other optimization parameters are defined as architecture parameters. These transfer functions include identity function (linearity function), unipolar sigmoid function (logistic function), bipolar sigmoid function (hyperbolic tangent function) and radial basis function (Gaussian function). Therefore, the j^{th} node in the i^{th} layer, $h_{i,j}$, can be described by the following equations.

1) Identity function:

$$h_{i,j} = s_{i,j} \quad i = 1, 2, \ldots, L+1; j = 1, 2, \ldots, n_i \tag{1}$$

where $$s_{i,j} = \sum_{k=0}^{n_{i-1}} w_{i,j,k} h_{i-1,k} - b_{i,j} \tag{2}$$

$$W_{i,j} = [w_{i,j,1}, w_{i,j,2}, \ldots, w_{i,j,n_{i-1}}, b_{i,j}] \tag{3}$$

$$H_{i-1} = [h_{i-1,1}, h_{i-1,2}, \ldots, h_{i-1,n_{i-1}}, -1]^T \tag{4}$$

In above equations, $w_{i,j,k}$ stands for a connection weight, $b_{i,j}$ for a node bias, $h_{0,k}$ for an input node and $h_{L+1,k}$ for an output node.

2) Unipolar sigmoid function:

$$h_{i,j} = \frac{1}{1+\exp(-s_{i,j})} \qquad i = 1,2,\ldots,L+1; \quad j = 1,2,\ldots,n_i \tag{5}$$

where $\exp(\cdot)$ is an exponential function and the value of $h_{i,j}$ is in the range (0, 1).

3) Bipolar sigmoid function:

$$h_{i,j} = \frac{1+\exp(-s_{i,j})}{1+\exp(-s_{i,j})} \qquad i = 1,2,\ldots,L+1; \quad j = 1,2,\ldots,n_i \tag{6}$$

where the value of $h_{i,j}$ is in the range (−1, 1).

4) Radial basis function:

$$
\begin{aligned}
h_{i,j} &= \exp\left(-\sum_{k=0}^{n_{i-1}} \frac{(h_{i-1,k} - w_{i,j,k})^2}{2b_{i,j}^2}\right) \\
&= \exp\left(-\frac{\left\|H_{i-1}^{*} - W_{i,j}^{*}\right\|^2}{2b_{i,j}^2}\right) \quad i = 1,2,\ldots,L+1; \quad j = 1,2,\ldots,n_i
\end{aligned}
\tag{7}
$$

where $W_{i,j}^{*} = [w_{i,j,1}, w_{i,j,2}, \ldots, w_{i,j,n_{i-1}}]$ (8)

$$H_{i\text{-}1}^{*} = [h_{i-1,1}, h_{i-1,2}, \ldots, h_{i-1,n_{i-1}}]^T \tag{9}$$

The value of radial basis function is in the range (0, 1).

In Figure 1, three sets of node genes are created to correspond to two possible hidden layers and one output layer. Each set of node genes is composed of three integer-valued codes, respectively corresponding to three possible hidden nodes in the first hidden layer, three possible hidden nodes in the second hidden layer and three output nodes in the output layer. The meanings of node genes are defined as follows: “0” denoting an inactive (nonexistent) node, “1” an identity function, “2” a unipolar sigmoid function, “3” a bipolar sigmoid function, and “4” a radial basis function. In Figure 1, the node genes of the second set indicate that the second node in the second hidden layer is inactivated, the first and the third nodes in the second hidden layer are activated, and the corresponding active transfer functions are bipolar sigmoid function and unipolar sigmoid function. In addition, the node genes of the third set are 1, 4 and 1. It means that in the output layer two identity functions and one radial basis function are chosen to donate the network outputs.

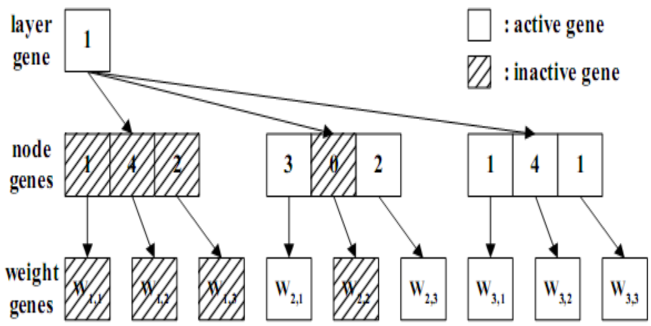

Figure 1: Genotype chromosome of a neural network

The MIHDE algorithm is successfully applied to many complex mixed-integer optimization problems [13-15]. It is stated in [15] in detail. Therefore, the MIHDE algorithm is used to conduct the training process in this paper. To evaluate an ANN, the objective function may be defined according to the performance requirements. For simplicity, the mean squared error function (MSE) is used and stated by the following equation:

$$E = \frac{1}{2mn_{L+1}} \sum_{k=1}^{m} \sum_{j=1}^{n_{L+1}} (d_j(k) - o_j(k))^2 \quad (10)$$

where o_j is an output of the neural network, d_j is the corresponding desired output, and m is the number of training data patterns. It is important to note that equation (10) is differentiable with respect to the real-valued weight parameters, but it is non-differentiable with respect to the integer-valued architecture parameters. Thus, the gradient based training methods, as back propagation algorithm, cannot be applied here to determine the optimal structure of a neural network. In contrast, this demonstrates that the MIHDE algorithm is suitable for the optimization design of ANNs.

Artificial Neural Network In Cancer Prediction

ANN is a branch of computational intelligence that employs a variety of optimization tools to learn from past experiences and use this prior training to predict and identify new patterns. In this neural network models have been used for the prediction of Colon Rectum Cancer.

Table 1: Network Information

Input Layer	Covariates	1	smoking
		2	obesity
		3	redmeat
		4	phy
		Number of Units[a]	5
		Rescaling Method for Covariates	Standardized
Hidden Layer(s)		Number of Hidden Layers	1
		Number of Units in Hidden Layer 1[a]	2
		Activation Function	Hyperbolic tangent
Output Layer	Dependent Variables	1	deathrate
	Number of Units		1
	Rescaling Method for Scale Dependents		Standardized
	Activation Function		Identity
	Error Function		Sum of Squares

Table 2:

Training	Sum of Squares Error	.041
	Relative Error	.014
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:00.000
Testing	Sum of Squres Error	.133
	Relative Error	0.84

Dependent Variable: deathrate

a. Error computations are based on the testing sample

The models were simulated using a variety of parameters and they were tested using many combination of parameters in independent experiments. The optimal prediction data for various ANN models were obtained by comparing with the parameter of error estimates such as Mean Square Error (MSE), Mean Absolute Percentage Error (MAPE).

Table 3:

	Importance	Normalized Importance
smoking	.159	46.4%
obesity	.344	100.0%
redmeat	.329	95.7%
phy	.046	13.4%
medicines	.122	35.4%

Computational Example

A benchmark problem is used to test the performance of the MIHDE algorithm in neural network design. For implementation, the setting parameters used in MIHDE are listed as follows: population size $N_P = 500$, crossover factor $\rho_c = 0.5$, and two tolerances $\varepsilon_1 = \varepsilon_2 = 0.1$. In addition, the neural network is assumed to have at most 2 hidden layers and at most 5 hidden nodes in each layer.

Experimental Results of Multi Layer Perceptron (MLP) Neutral Network

The feed forward neural network architecture used in this experiment consists of two hidden layer along with one input and output layer respectively. The transfer function in hidden layer neurons and output layer neurons are hyperbolic tangent and identity. The performance function used was MSE.

- x : p X_1 input vector
- h : Weighted sum of input stimuli
- v : m X_1 output vector of hidden layer
- g : Weighted sum of v_j
- y : n X_1 output vector of output layer
- w_{ij} : weight connecting i^{th} unit of output layer and j^{th} unit of hidden layer
- w_{jk} : Weight connecting j^{th} unit of hidden layer to k^{th} unit of input layer
- y : Actual output
- y^d : Desired output.

where i, j and k indices referring to the neurons belonging to the output, hidden and input layers respectively, p, m and k are the number of neurons in input, hidden and output layer respectively.

Table 4: Case Processing Summary

		N	Percent
Sample	Training	7	63.6%
	Testing	4	36.4%
Valid		11	100.0%
Excluded		0	
Total		11	

Predictors for the ANN : There are 5 predictors for ANN which re smoking, obesity, red meat eating, physical activity and usage of aspirin and other medicines respectively.

Conclusions

In this paper, a Mixed-Integer encoding scheme is proposed to represent the ANN's architecture and weight parameters including types of transfer functions, network topology and connection weights. And then a mixed-integer evolutionary algorithm, MIHDE, is used to optimize the weights and architectures of ANNs. Finally, a benchmark problem is used to test the capability of the optimized ANN by the MIHDE algorithm. Computational results demonstrate that the optimized ANN model can effectively predict the Mackey-Glass colon rectum cancer. The successful results show a significant progress in research on the optimization of ANNs. The non-linearity component of the relationship can be successfully dealt with using ANN. This shows the effectiveness of ANN in the prediction of colon rectum cancer.

References

[1] X. Yao, "Evolving artificial neural networks," Proceedings of the IEEE, vol. 87, no. 9, pp. 1423–1447, 1999.

[2] P.J. Angeline, G.M. Sauders, and J.B. Pollack, "An evolutionary algorithm that constructs recurrent neural networks," IEEE Trans. Neural Networks, vol. 5, pp. 54–65, 1994.

[3] J.R. Koza and J.P. Rice, "Genetic generation of both the weights and architecture for a neural network," in Proc. IEEE Int. Joint Conf. Neural Networks, vol. 2, pp. 397–404, 1991.

[4] J.R. Koza and J.P. Rice, "Genetic generation of both the weights and architecture for a neural network," in Proc. IEEE Int. Joint Conf. Neural Networks, vol. 2, pp. 397–404, 1991.

[5] K.S. Tang, C.Y. Chan, K.F. Man, and S. Kwong, "Genetic structure for NN topology and weights optimization," in Proc. 1st IEE/IEEE Int. Conf. Genetic Algorithms in Engineering Systems: Innovations and Applications (GALESIA'95), pp. 250–255, 1995.

[6] M. Mandischer, "Evolving recurrent neural networks with non-binary encoding," in Proc. IEEE Int. Joint Conf. Evolutionary Computation, vol. 2, pp. 584–589, 1995.

[7] G.F. Miller, P.M. Todd and S.U. Hegde, "Designing neural networks using genetic algorithms," in Proc. 3rd Int. Conf. Genetic Algorithms and Their Applications, pp. 379–384, 1989.

[8] Y.C. Lin, K.S. Hwang and F.S. Wang, "Co-evolutionary hybrid differential evolution for mixed-integer optimization problems," Engineering Optimization, vol. 33, no. 6, pp. 663–682, 2001.

[9] Y.C. Lin, K.S. Hwang and F.S. Wang, "Co-evolutionary hybrid differential evolution for mixed-integer optimization problems," Engineering Optimization, vol. 33, no. 6, pp. 663–682, 2001.

[10] Y.C. Lin, K.S. Hwang, and F.S. Wang, "A mixed-coding scheme of evolutionary algorithms to solve mixed-integer nonlinear programming problems," Computers and Mathematics with Applications, vol. 47, pp. 1295–1307, 2004.

[11] M.C. Mackey and L. Glass, "Oscillation and chaos in physiological control systems," Science, vol. 197, pp. 287–289, 1977.

[12] M.G. Kendall and A. Stuart, (1966), "The advanced theory of statistics", Vol. 3. Design and Analysis and Time-Series, Charles Griffin & Co. Ltd., London, United Kingdom.

[13] G.E.P. Box and G.M. Jenkins (1970), "Time series analysis: Forecasting and control", San Francisco: Holden-Day.

[14] C.W.J. Granger and P. Newbold, (1986), "Forecasting Economic Time Series" (Academic Press, San Diego).

[15] G.E.P. Box, G.M. Jenkins and G.C. Riesel, (1994), "Time Series: Analysis: Forecasting and control", Pearson Education, Delhi.

[16] P.J. Brockwell and R.A. Davis, (1996), “Introduction to time series and forecasting” Springer.
[17] David B. Fogel, Eugene C. Wasson, Edward M. Boughton, Vincent W. Porto, and Peter J. Angeline, “Linear and Neural Models for Classifying Breast Masses”, IEEE transactions on medical imaging, vol. 17, no. 3, June 1998, pp 485–488.
[18] Heng-Da Cheng, Yui Man Lui, and Rita I. Freimanis “A Novel Approach to Microcalcification Detection Using Fuzzy Logic Technique”, IEEE transactions on medical imaging, vol. 17, no. 3, June 1998, pp 442–450.
[19] Ky Van Ha, “Hierarchical Radial Basis Function Networks”, 1998, EEE 1893 PP 1893–1898.
[20] S. Makridakis, S.C. Wheelwright and R.J. Hyndman, (1998), “Forecasting: methods and applications”, New York: John Wiley & Sons.