JOURNAL OF COMPUTING TECHNOLOGIES
ISSN 2278 – 3814

Available online at www.jctjournals.com
Volume 1, Issue 3, July 2012

MIXED GENETIC ALGORITHM APPROACH FOR FUZZY LOGIC CONTROL

SIVAKUMAR R[1]
Dr.M.SRIDHAR[2]
Research Scholar Dept of ECE BHARATH UNIVERSITY India[1]
Dept of ECE BHARATH UNIVERSITY India[2]
siva999.siva999@gmail.com

Abstract

Most of the real world problems in engineering, medicine, industry, science and business involves data classification task. Data classification takes labeled data samples and generates a classifier model that classifies new data samples into different predefined groups or classes. This thesis explores the application of Computational Intelligent Techniques like Artificial Neural Network (ANN) and Fuzzy Logic (FL) for solving data classification problems. Artificial Neural Network has emerged as an important tool for solving data classification problem. Artificial Neural Network is an information-processing paradigm inspired by the way the brain process information. It is composed of a large number of highly interconnected processing elements (neurons) working in parallel to solve a specific problem. Due to their powerful non-linear function approximation and adaptive learning capabilities ANN can make accurate data classification.

Keyword: iris Data, Wine Data, *TCPDUMP Data*

Introduction

When designing a fuzzy system using Genetic Algorithms, the first important consideration is the representation strategy to be followed. Binary strings are used to represent the rule set. In the rule set each variable is represented by a two bit substring. An additional two bit is used to represent the output class. In order to design a compact rule set only ‘nR’ rules are represented in the genetic population. Within that ‘nR’ rules, to select the optimal rules, a rule selection bit is used. The membership function is represented as floating point numbers. Each input variable is fuzzified into three linguistic values (low, medium and high). Trapezoidal membership function is used to represent the lower and higher values of the input variable and traiangular membership function is used to represent the medium value of the input variable. With this representation a typical chromosome will look like the following:

0 01 10 ... 00 01 1 10 11 … 10 00
…. 0 01 00… 11 10

RS IP_1 IP_2 IP_n OP RS IP_1 IP_2 IP_n OP
RS IP_1 IP_2 IP_n OP

Rule 1 **Rule 2**
Rule nR

5.6,6.1,7.0 2.6,3.2,3.8 2.4,3.9,5.4 ………
0.3,0.6,0.9 1.2,1.5,1.8 2.4,2.7,3.1

MF_{low} MF_{medium} MF_{high}
MF_{low} MF_{medium} MF_{high}

Input1 **Input n**

This type of representation has a number of advantages over binary coding for membership function. The efficiency of the GA is increased as there is no need to convert the input variables to the binary type.

Iris Data Classification

Iris data set consists of 150 four dimensional vectors representing 50 plants each of three species *iris setosa, iris versicolor and iris virginica.* The four input features are sepal length, sepal width, petal length and petal width. All these input features are continuous variables. Each input variable is represented by three fuzzy sets and a total of 7 points are needed to represent each variable. The range of each membership function point is computed dynamically.

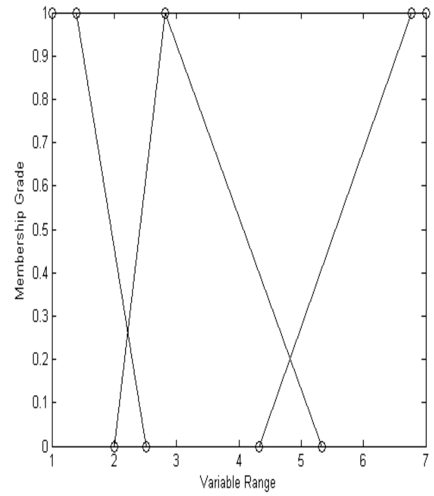

Figure 1 Optimal Membership function for the variable petal length

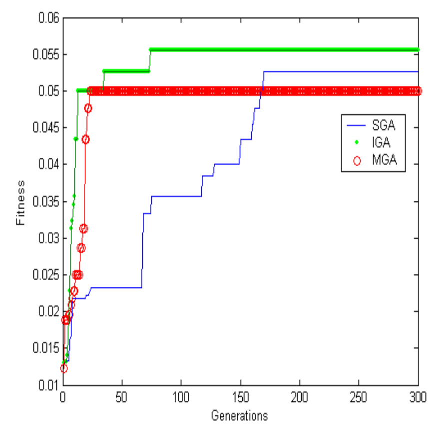

Fig : 2 Convergence behavior of MGA compared with SGA and IGA

The convergence behavior of the proposed MGA is compared with SGA and IGA are shown in Figure 1.

Wine Data Classification

Next the performance of the proposed approach is tested on Wine data. The wine data contains the chemical analysis of 178 wines grown in the same region in Italy but derived from three cultivators which is designated as 3 output classes. The 13 attributes available for classification are: alcohol, malic acid, ash, alkalinity of ash, magnesium, total phenols, flavanoids, nonflavanoids phenols, proanthocyaninsm color intensity, hue, OD280/OD315 of diluted wines and praline. Among the 13 input features 11 features are continuous and 2 are discrete. The same representation strategy that is followed in Iris data set is followed here also. A total of 91(13×7) points are needed to represent all the membership functions of all the input variables. A maximum of forty rules were generated initially. Each rule needs 29 bit binary string and hence a total of 1160(29×40) binary bits are used to represent the entire rule set.

TCPDUMP Data Classification

Finally, the proposed approach was applied to design the classifier for the TCPdump data. The TCPdump data set is actually a Windows NT attack data set collected from the 2000 DARPA Intrusion Detection Specific Scenario Data sets. It consists of a total of 294 data with 259 normal cases and 35 abnormal cases with five input features (Source IP address, Destination IP address, Source Port Number, Destination Port Number and the Protocol used) and two output classes (Normal, Abnormal). All the input features are discrete in nature.

Mixed Genetic Algorithm

Starting with an initial population, the genetic algorithm exploits the information contained in the present population and explores new individuals by generating offspring using the three genetic operators namely, reproduction, crossover and mutation which can then replace members of the old generation. Fitter chromosomes have higher probabilities of being selected for the next generation. After several generations, the algorithm converges to the best chromosome, which hopefully represents the optimum or near optimal solution. The above process is pictorially represented.

Conclusion

Mixed Genetic Algorithm (MGA) approach for designing a fuzzy classifier which improves the convergence speed and quality of the solution. In the proposed MGA, floating point numbers are used to represent the membership function and binary strings are used to represent the rule set. This type of representation has a number of advantages over binary coding for membership function. The efficiency of the GA is increased as there is no need to convert the input variables to the binary type. For effective genetic operation, modified form of crossover and mutation operators which can deal with the mixed string are proposed.

In a standard Simple Genetic Algorithm (SGA), crossover is the main genetic operator responsible for the exploitation of information while mutation brings new nonexistent bit structures. It is widely recognized that the SGA scheme is capable of locating the neighborhood of the optimal or near-optimal solutions, but, in general, SGA requires a large number of generations to converge. To overcome this problem, this work proposes an Improved Genetic Algorithm (IGA) approach which incorporates a set of advanced and problem-specific genetic operators (Durairaj *et al*, 2006) in addition to the basic cross over and mutation operator applied in SGA in order to improve its convergence and the quality of the solutions. The following problem-specific operators have been used in this work.

References

[1]. Cordon O., Gomide F. and Herrera F. (2004) 'Ten Years of Genetic Fuzzy Systems: Current Framework and New Trends', Fuzzy Sets and Systems, Vol.141, No.1, pp. 5-31.

[2]. Dasarathy B.V. (1991) 'Nearest-Neighbor Classification Techniques', IEEE Computer Society Press, Los Alomitos, CA.

[3]. Demsar J. (2006) 'Statistical Comparisons of Classifiers over Multiple Data Sets', J. of Machine Learning Research, Vol. 7, pp.1-30.

[4]. Denoeux T. (1995) 'A k-nearest neighbor classification rule based on Dempster- Shafer theory', IEEE Trans. Syst., Man, Cybern., Vol. 25, pp. 804–813.

[5]. Devaraj D. and Yegnanarayana B. (2005) 'Genetic Algorithm-Based Optimal Power Flow for Security Enhancement', IEE Proc. on Generation, Transmission and Distribution, Vol. 152, No. 6, pp. 899 – 905.

[6]. Devaraj D., Yegnanarayana B. and Ramar K. (2002) 'Radial Basis function networks for fast contingency ranking', Int. J. of Electrical Power and Energy System, Vol. 24, No. 5, pp.387-395.

[7]. Dubois D., Prade H. and Testermale C. (1988) 'Weighted fuzzy pattern matching', Fuzzy Sets System, Vol.28, No.3, pp.313-331.

[8]. Durairaj S., Devaraj D. and Kannan P.S (2006), 'Voltage stability constrained reactive power planning using Improved Genetic Algorithm', Int.J. of Water and Energy Vol. 1, pp.56-64.

[9]. Edgoose T. and Allison L. (1999) 'MML Markov classification of sequential data', Stats. and Comp., Vol. 9 , No. 4, pp.269-278.

[10]. Friedman J.H. (1987) 'Exploratory Projection Pursuit', J. of American Statistical Association, Vol. 32, No. 397.

[11]. Ghosh J. and Chakravarthy S. V. (1994) 'Rapid kernel classifier: A link between the self organizing feature map and the radial basis function network', Journal of Intelligent Material Systems and Structures, Vol.5, pp.211-219.

[12]. Glorfeld L.W. (1996) 'A methodology for simplification and interpretation of back propagation-based neural networks models', Expert Syst. Application, Vol.10, pp.37-54.

[13]. Goldberg D.E. (1989) 'Genetic Algorithms in Search, Optimization, and Machine', Addison-Wesley.