Analysis of Fuzzy Rule Optimization Models

Amudha J[#1], D Radha[#2], Smitha S[#3]
[#]Department of Computer Science and Engineering
Amrita School of Engineering, Amrita Vishwa Vidyapeetham,
Bangalore, Karnataka, India.
[1]j_amudha@blr.amrita.edu,
[2]d_radha@blr.amrita.edu,
[3]smitha2112@gmail.com

Abstract— Optimization without losing the accuracy and interpretability of rules is a major concern in rule based system. Fuzzy Inference system characterized by uncertainty tolerance is the best way to represent a knowledge based system. Optimization of rule based systems starts by incorporating self-learning ability to a fuzzy inference system. This can be achieved by neural networks, there by developing a neuro fuzzy inference system. This paper analyses different neuro fuzzy inference systems .The analysis has been performed in different types of datasets in terms of dimensionality and noises. Analysis results concludes that the neuro fuzzy model DENFIS (Dynamically Evolving Neuro Fuzzy Inference System) shows an improved performance when handling with high dimensional data. Simulation results on low dimensional data exhibits similar performance in ANFIS (Adaptive Neuro Fuzzy Inference System) and Denfis.

Keyword-Computational intelligence, rule optimization, Anfis, Denfis, neuro-fuzzy.

I. INTRODUCTION

The computational Intelligence paradigm (CI) inspired by nature, comprises of Artificial Neural Network (ANN), Evolutionary Algorithm (EA), Swarm Intelligence (SI), Fuzzy Logic (FL) etc. These are different flavours for increasing efficiency in terms of memory and computational power which alternatively affects other issues towards perfection i.e. convenience and intelligence. Hybrid and non-hybrid algorithms optimizes the issues through CI paradigms. Heuristic problem solving approach which emphasizes in solving a problem by formulating a set of rules or a set of procedures have been widely accepted by different realms of technology. Rule based systems falls under the category of such heuristic approach of problem solving. Rule based expert system or knowledge based systems can solve the problem based on the rules, which are formulated through a previously represented knowledge. Rule based systems represents human knowledge in the form of easily interpretable IF…..THEN rules.

Fuzzy Inference system are rule based systems which deals with imprecise data. In fuzzy inference system[1] the input variables are represented using linguistic terms, since the fuzzy logic deals with uncertainty .The fuzzy inference system is composed of a knowledge base, Inference engine, Fuzzifier and Defuzzifier. Fuzzifier maps the input variable to its membership function, and defuzzifier convert the output membership functions to crisp value. Since Fuzzy logic deals with imprecise data it has made a great impact in different fields of technology like decision making, classification, automatic control of machines, computer vision etc. The main disadvantage of fuzzy inference system is that human intervention is required for formulating the fuzzy rules. Rule generation requires an expertise knowledge in that area and the process is an exhaustive task in case of complex application. This disadvantage leads the direction of exploration towards the development of fuzzy logic systems that have the ability to learn from experience. The best computational intelligent technique which can be incorporated with fuzzy inference system for self-learning are neural networks.

Neural networks, a computational intelligence technique is a connectionist system of nodes in different layers. Neural network is mimicking the learning ability of human brain system. The integrated model neuro-fuzzy network [2] solves the problem of additional overhead in learning and generating the nodes. The further research was in the field of optimizing the rules of Fuzzy Inference System.

The rules of fuzzy inference system is a combination of different inputs, therefore while optimizing the rules of a fuzzy inference system number of inputs also have an important rule. This paper compares two neuro fuzzy models Anfis [3] and Denfis [4] and analyses the impact of different datasets.

II. RELATED WORK

In the literature, majority of the works were considering input selection and rule selection separately. Input selection can be performed in Fuzzy Inference system. The strength of neuro fuzzy network lies in the training capability. The most widely used training technique in the literature is ANFIS. Based on the input space partitioning Anfis are of two types, Anfis-Grid and Anfis-Sub. Anfis-Grid follows grid partition technique and Anfis-Sub follows sub-clustering technique of input space partitioning. Grid partition considers all the combination of inputs while constructing the rules, so this is a good approach when input attributes are less. Anfis-Sub was the solution when no.of inputs is beyond a certain limit. Evolving Fuzzy neural network (EFuNN) [5] is a neuro-fuzzy network which provides online learning. It learns faster and new connections and new neurons are created during the operation of the system. Denfis have both online and offline method of learning the online method of learning. It uses evolving clustering method (ECM) for online learning and constrained Evolving clustering method (ECMc) for offline learning as the input space partitioning technique.

TABLE I. Comparison of Different Rule Based Approaches

Model	Learning ability	Automatically generate membership function	Learning rate	Online model	Interpretability of rules	Curse of dimensionality	Support Multiple output models
Fuzzy Inference system[1]	no	no	Low	no	High	Yes	Yes
Neuro Fuzzy Inference System [2]	Yes	Yes	High	no	High	Yes	Yes
Anfis [3]	Yes	Yes	High	no	High	Yes	no
EFuNN [5]	Yes	Yes	High	Yes	Low	no	Yes
Denfis [4]	Yes	Yes	High	Yes	Low	no	Yes
VSANF [11]	Yes	Yes	Low	Yes	Low	no	Yes
TWNFI [12]	Yes	Yes	Low	Yes	Low	no	Yes

Various input selection models in literature can be achieved in a variety of ways: using the regularity criterion [6], the geometric criterion [7], individual discrimination power [8], and entropy variation index [9] or by using Gram–Schmidt orthogonal least squares [10].Variable selection algorithm for the construction of MIMO operating point dependent neurofuzzy networks (VSANF) [11] is a multiple input multiple output model, which tackles the curse of dimensionality. A transductive neuro-fuzzy inference system with weighted data normalization for personalized modeling (TWNFI) [12] depicts the most significant input variables which requires more training time which can be a disadvantage of this model. The advantages and disadvantages of various model described in the literature survey are summarised in table 1.Different models are compared with various parameters like learning ability, automatic generation of membership functions, learning rate, online learning ability ,rule interpretability, curse of dimensionality and multiple output models. Learning ability indicates the adaptability of the model and the faster learning rate corresponds to the reduced training time .Dimensionality refers to the number of features used to represent the dataset. Due to memory constraints high dimensional data are not equally accepted by all models.

III. ARCHITECTURE

Anfis described in the literature is the widely used neuro fuzzy model owing to its faster learning rate and rule interpretability. In this paper Anfis is compared with Denfis model, which also uses clustering method for input space partitioning .The basic architecture of Anfis and Denfis is shown in Fig.1

A. Anfis Architecture

Adaptive neuro fuzzy inference system (ANFIS) was developed by Jang. Anfis uses Takagi Sugeno type inference system, where output is either a crisp constant or a function of linear or non-linear system. Anfis is a five layered structure as shown in Fig.1.A, B represents the inputs given to the model. The functionalities of each of the layers are described below.

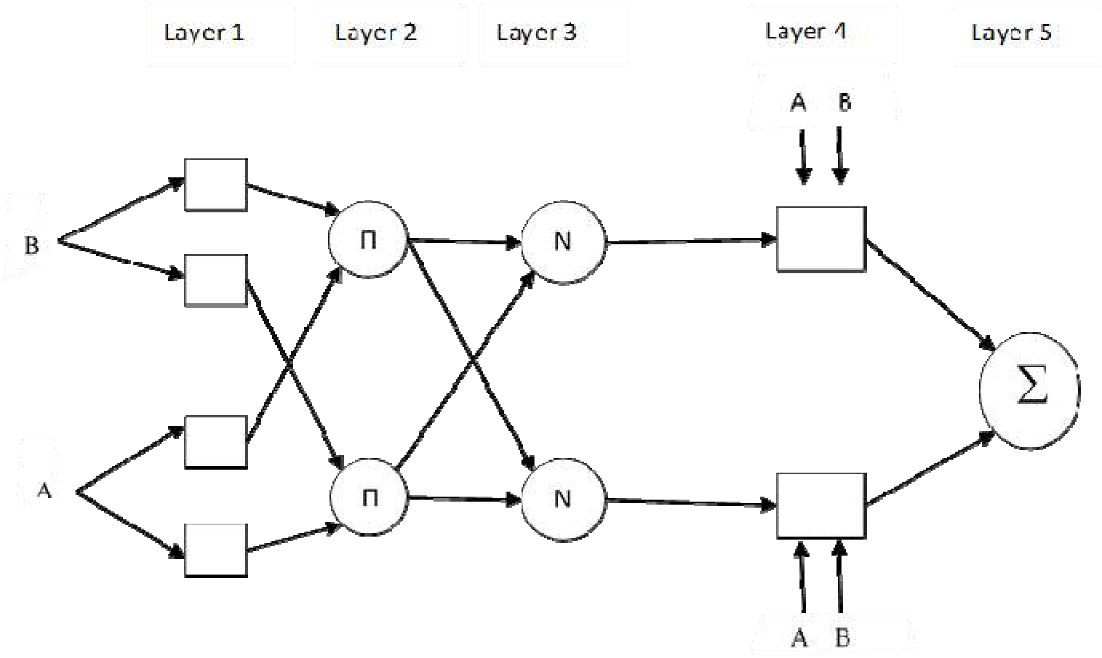

Fig. 1. Basic Architecture of Anfis and Denfis(from DENFIS,2000)

Layer 1: This layer provides the input in the desired format for the appropriate membership function and calculates the degree of membership for each of the nodes. The membership function can be gaussian, triangular, trapezoidal etc.

Layer 2: The output of the nodes in this layer is calculated as the product (Π) of the corresponding nodes from all the inputs in the previous layer

Layer 3: The output of the previous layer is subjected to normalisation (N) in this layer.

Layer 4: The output of this layer is a linear function of all the initial inputs and normalised value from the previous layer

Layer 5: This layer produces an output which is a summation (Σ) of inputs received across all the nodes from the previous layer.

B. Denfis Architecture

In Denfis fuzzy inference rules are created based on scatter partitioning of input space variables. Evolving Clustering method has been used to cluster the input space. Each input vector belongs to one or more clusters. The degree of membership of each input vector in its cluster is find out considering the membership function as a triangular membership function

Layer 1: This layer performs the clustering of input vectors and determines the cluster centre. Cluster centre and the distance threshold determines the three parameters for the calculation of membership degrees. The general form of a triangular membership function is given in (1),where x is the input to be fuzzified a, b, c are the parameters of triangular membership function , b is the centre for cluster along x dimension, $a = b - d \times dthr$ and $c = b + d \times dthr$,d varies from 1.2 – 2 and dthr is the threshold value of the clustering parameter.

$$\mu(x) = mf(x,a,b,c) = \begin{Bmatrix} 0, x \le a \\ \dfrac{x-a}{b-a}, a \le x \le b \\ \dfrac{c-x}{c-b}, b \le x \le c \\ 0, c \le x \end{Bmatrix} \quad (1)$$

Layer 2: The output of the nodes in this layer is calculated as the product (Π) of the membership degrees of input variables in the previous layer.

Layer 3: Layer 2 output is subjected to normalisation (N) in this phase.

Layer 4: This layer calculates each rule output. In Denfis first order Takagi Sugeno kang (TSK) [11] type inference systems are used. Each rule output of the TSK is calculated using (2), where β's are the rule consequent and is calculated using the Least Square Estimator LSE [11].

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ...\beta_n x_n \tag{2}$$

Layer 5: This layer calculates the final output, which is the weighted sum of each rule's output.

IV. Simulation and analysis of results

Simulation and analysis for Anfis and Denfis model was done using datasets mentioned in table 2. The details of the datasets used for training and testing purpose have been mentioned in table.2 .The firsts three datasets have been sourced from the UCI repository [13], the last one is a target search image dataset which is extracted using the Vocus [14] model and have 13 attributes representing intensity, orientation and color of the image. This image dataset is characterized by the presence of noisy values. These two models have been compared with the number of rules generated and the average testing error obtained when the testing data is applied on the generated rules.

A. *Analysis of Anfis and Denfis*

According to the FIS proposed in [15], the number of fuzzy rules for Iris dataset classification with three membership function each, having 16 rules with an accuracy of 98.6%. Fig 2 shows the results of Anfis and Denfis for the iris datasets. The optimal point is obtained when the number of rules are 5 with an average testing error of 0.15.This scenario corresponds to choosing the clustering threshold value which represents the cluster radius as 1.25 and 0.2 in Anfis and Denfis respectively. For all other points the error is as shown in Fig .2. Simulation results of Denfis on iris dataset the average testing error hits the minimum value when the number of rules are 4. This scenario corresponds to choosing the clustering threshold value which represents the cluster radius as 1.25 and 0.2 in Anfis and Denfis respectively. Since iris dataset has only 4 attributes to describe them, both Anfis and Denfis does not show much variation. In order to analyse the impact of input on rule generation, datasets having more number of attributes have to be considered. Wine dataset contains 13 distinct attributes to classify the data into three classes. Anfis and Denfis generates 45 and 8 optimal rules respectively for the wine dataset as shown in fig.3. This optimal value is obtained when the cluster parameter is 2 in Anfis and .05 in Denfis .So when the number of attributes is more Denfis shows remarkable improvement in performance. In order to ascertain this observation datasets with large no.of attributes are also considered. Ionosphere dataset is selected for this purpose which have 34 attributes .The same procedure is repeated for ionosphere dataset. The clustering threshold value which gives this result is 2.25 for Anfis and.25 for Denfis The results were enough to strengthen the inference made by the previous analysis since the no.of rules has been reduced from 90 to 25 as in Fig.4..The analysis results of Anfis and Denfis on the noisy image dataset the optimal rules obtained in Denfis is 12 and in Anfis is 18 as shown in Fig.5. The optimal clustering parameter values are 0.4 in Anfis and 0.92 in Denfis. The comparative study of the two models Anfis and Denfis reveals that the number of input have an important role in rule optimization process. The Denfis model is the better approach when dealing with high dimensional data, whereas in case of low dimensional data Anfis and Denfis have similar performance.

In addition to the improved performance of Denfis with high dimensional data, Denfis takes considerably less time for training the input samples, prior to the rule generation phase. The ECM algorithm used for the input space partitioning, which is an online single pass algorithm, owes to the reduced time complexity of Denfis model. Similar training time (in the order of microseconds) is observed with Anfis and Denfis for low dimensional data. But for training high dimensional data a significant difference in training time is observed in Anfis (order of seconds) and Denfis (order of milliseconds).

TABLE 2. Dataset Description

Dataset name	**No.of attributes**	**No.of instances**	**No.of training data / No.of testing data**
Iris dataset	4	150	90 / 60
Wine Dataset	13	178	100 / 78
Ionosphere	34	351	200 / 151
Image Dataset	13	2500	1500 / 1000

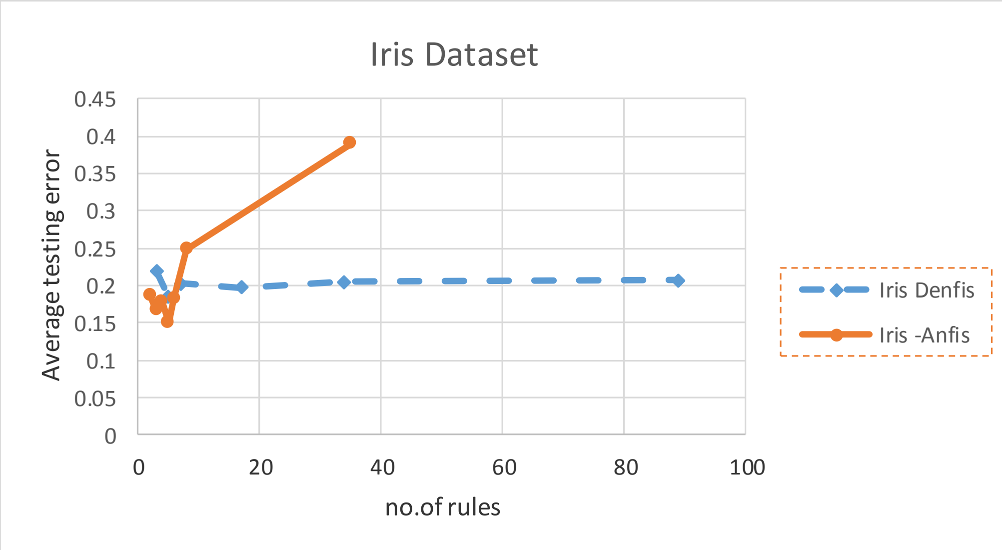

Fig. 2. Analysis result of Anfis and Denfis on Iris Dataset

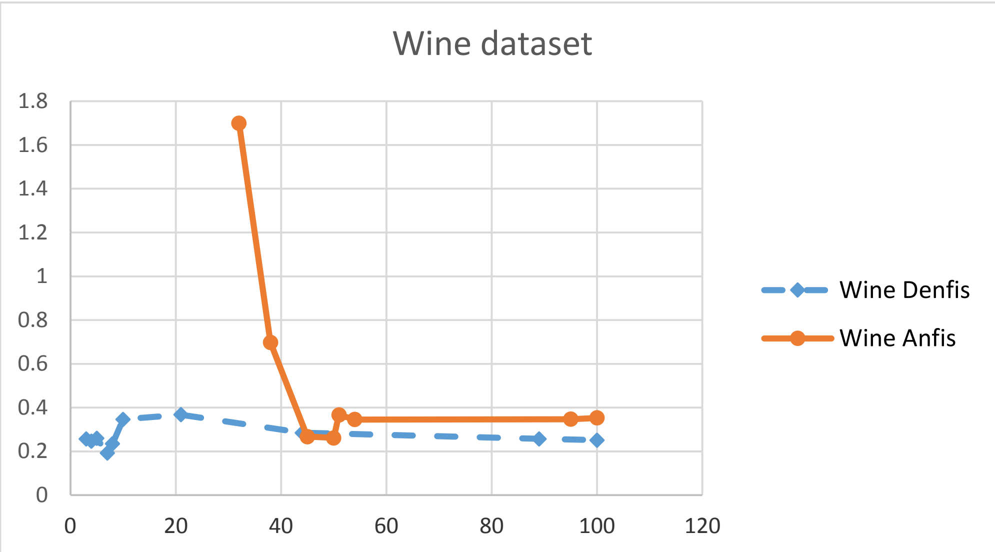

Fig. 3. Analysis result of Anfis and Denfis on Wine Dataset

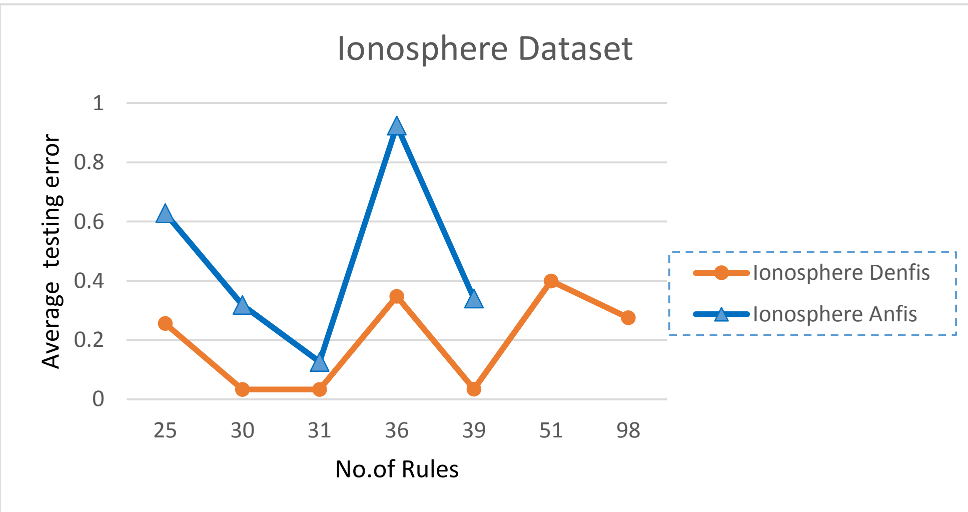

Fig.4.Analysis results of Anfis and Denfis on Ionosphere Dataset

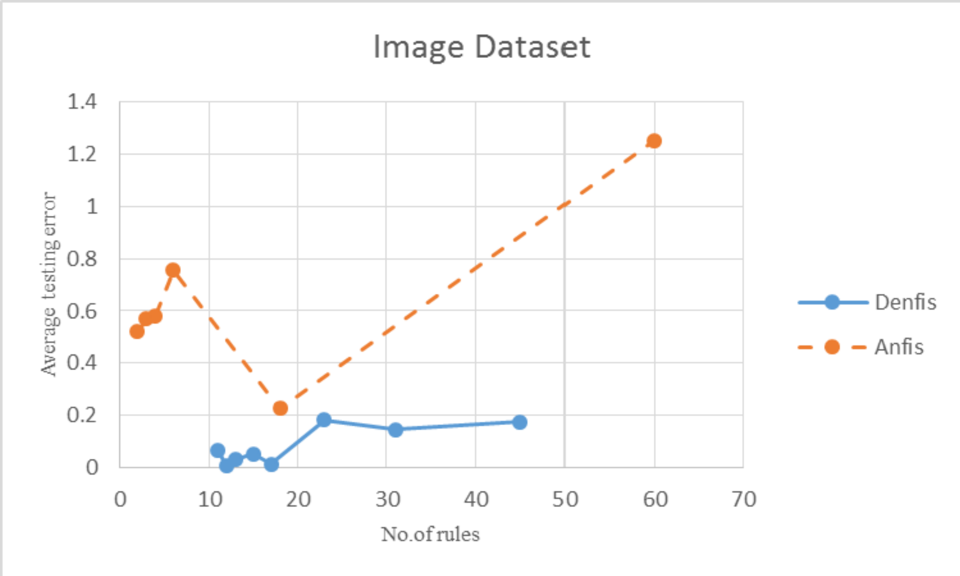

Fig.5.Anaysis result of Anfis and Denfis on Image Dataset

V. CONCLUSION

Fuzzy rule based system is a renowned method to deal with imprecise data. Incorporating neural network to the fuzzy rule based systems makes the system an adaptive model. Anfis and Denfis belongs to this category of adaptive neuro fuzzy systems, which vary in the learning methods which are used for implementation. A comparative study of Anfis and Denfis models to analyze their impact on rule generation has been performed. The result on classification datasets show that Denfis model is better than Anfis when dealing with high dimensional data, though performances are comparatively similar on small dimensions. The work arrives at the conclusion that number of rules is dependent on the number of input attributes, hence a reduced rule base is obtained by selecting the relevant attributes which can successfully identify the data element. Rule optimization is to a large extent influenced by the clustering techniques used in these neuro fuzzy models. The improved performance of Denfis model for rule optimization owes to the evolving cluster method implemented. ECM enhances the performance by identifying broad visualizations of similar data points into the same cluster and hence capable of optimizing rules.

ACKNOWLEDGMENT

Iris, Wine and Ionosphere datasets used in this work has been obtained from the UCI Repository of Machine Learning Databases. We would like to thank UCI Repository and its contributors for providing the datasets.

REFERENCES

[1] L. Wang and J. M. Mendel, “Fuzzy basis functions, universal approximation, and orthogonal least-squares learning,” IEEE Trans. Neural Netw., vol. 3, no. 5, pp. 807–814, Sep. 1992.

[2] Shin –ichi Horikawa, Takeshi Furuhashi and Yoshiki Uchikawa On Fuzzy Modelling using Fuzzy Neural Networks with the Back-Propagation Algorithm,’*IEEE Transactions on Neural Networks* ‘ VOL.3, No.5,September 1992.

[3] J.-S. R. Jang, “ANFIS: Adaptive-network-based fuzzy inference system,” *IEEE Trans. Syst. Man, Cybern.*, vol. 23, no. 3, pp. 665–685, May/Jun.1993.

[4] Q. Song and N. Kasabov, “Dynamic evolving neuro-fuzzy inference system (DENFIS): On-line learning and application for time-series prediction,” *IEEE Trans. Fuzzy Syst.*, vol. 10, no. 2, pp. 144–154, Apr. 2000.

[5] N. K. Kasabov, “On-line learning, reasoning, rule extraction and aggregation in locally optimized evolving fuzzy neural networks,” *Neurocomputing*, vol. 41, pp. 25–45, 2001.

[6] M. Sugeno and T.Yasukawa, “Afuzzy-logic-based approach to qualitative modelling,” *IEEE Trans. Fuzzy Syst.*, vol. 21, no. 1, pp. 7–31, Mar. 1993.

[7] M. R. Emami, I. B. Turksen, and A. A. Goldenberg, “Development of a systematic methodology of fuzzy logic modeling,” *IEEE Trans. Fuzzy Syst.*, vol. 6, no. 3, pp. 346–361, Aug. 1998.

[8] T.-P. Hong and J.-B. Chen, “Finding relevant attributes and membership functions,” *Fuzzy Sets Syst.*, vol. 103, no. 3, pp. 389–404, 1999.

[9] N. R. Pal, “Soft computing for feature analysis,” *Fuzzy Sets Syst.*, vol. 103, no. 2, pp. 201–221, 1999.

[10] S. Chen, S. A. Billings, and W. Luo, “Orthogonal least squares methods and their application to nonlinear system identification,” *Int. J. Contr.*,vol. 50, pp. 1873–1896, 1989.

[11] X. Hong and C. J. Harris, “Variable selection algorithm for the construction of MIMO operating point dependent neurofuzzy networks,” *IEEE Trans. Fuzzy Syst.*, vol. 9, no. 1, pp. 88–101, Feb. 2001.

[12] Q. Song and N.Kasabov, “TWNFI—A transductive neuro-fuzzy inference system with weighted data normalization for personalized modeling,”*Neural Netw.*, vol. 19, pp. 1591–1596, 2006.

[13] Lichman, M. (2013). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

[14] Simone Frintrop, “VOCUS: a visual attention system for object detection and goal-directed search,” Ph.D. dissertation, Rheinische Friedrich- Wilhelms-Universit¨at Bonn, Germany, July 2005, published 2006 in Lecture Notes in Artificial Intelligence (LNAI), Vol. 3899, Springer Verlag Berlin/Heidelberg.
[15] A. Sood, S. Aggarwal, Crossroads in classification: comparison and analysis offuzzy and neuro-fuzzy techniques, International Journal of Computer Applications 24 (2) (2011) 13–17.

AUTHOR PROFILE

Amudha J is an Associate Professor in Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bangalore- 560035. Her qualifications include B.E., Electrical and Electronics Engineering from Bharathiyar University, Coimbatore, India and M.E. Computer Science and Engineering, Anna University, Chennai, India and PhD from Amrita Vishwa Vidyapeetham, Coimbatore. Her areas of research include Image Processing, Computer Vision, Soft Computing, and Machine Learning. She has over 30 papers in national and international journals and proceedings. She has organized various technical workshops in the institution.

D Radha is an Assistant Professor in Amrita School of Engineering, Amrita Vishwa Vidyapeetham, and Bangalore-560035. Her Qualifications include B. E, Computer Science and Engineering from Madurai Kamaraj University, Madurai, India and M.Tech Computer Science and Engineering, Dr. MGR University, Chennai, India. Her areas of interest are Image Processing, Computer vision and High Performance computing.