

Usage of Simplified Fuzzy ARTMAP for improvement of classification performances

Emina Alickovic, Abdulhamit Subasi

International Burch University, Faculty of Engineering and Information Technologies, Francuske revolucije bb, Ilidža 71210 Sarajevo, Bosnia and Herzegovina

Article Info

Article history:

Received 17 Sep. 2013

Received in revised form 17 Oct 2013

Keywords:

Fuzzy ARTMAP, Simplified Fuzzy ARTMAP, classification, vigilance parameter

Abstract

This study presents a simplified fuzzy ARTMAP (SFAM) for different classification applications. The proposed SFAM model is synergy of fuzzy logic and adaptive resonance theory (ART) neural networks. SFAM is supervised network consisting of two layers (Fuzzy ART and Inter ART) that build constant classification groups in answer to series of input patterns. Fuzzy ART layer takes a series of input patterns and relate them to output classes. Inter ART layer functions in such a way that it raises the vigilance parameter of fuzzy ART layer. By combining this two layers, SFAM is capable to perform classification very efficiently and giving very high performances. Lastly, the SFAM model is applied to different simulations. The simulation results obtained for the three different datasets: Iris, Wisconsin breast cancer and wine dataset prove that SFAM model has better performance results than other models for these classification applications.

1. INTRODUCTION

One of the most frequent decision making problems is classification. A classification is the problem of prescribing of objects into predefined clusters or classes according to the amount of detected attributes associated with corresponding object, and offering several measures for deciding if a prescribed object belongs a specific group or not. Classification tasks are found in numerous decision making task in different areas such as medicine, science, industry etc. There are several approaches how to solve classification problems.

Statistical approach is traditional approach having a clear core probability modeling and this approach is based on the Bayesian decision theory [1]. Chief disadvantage of using statistical approaches is that they give good performances just in case primary assumptions are accurate. In order to obtain good performances, users should own an adequate understanding of both information characteristics and model abilities.

Another approach in solving classification problems is usage of different machine learning techniques. Different machine learning techniques have been proposed for these purposes. Techniques resulting in higher classification performances will provide more adequate evidence to classify objects to certain groups and to improve the classification accuracy. Support vector machines, k-NN, decision tree methods have been extensively used in classification problems. The aim of all these methods was to obtain the highest classification accuracy rate possible. In the recent years, special attention has also been given to neural networks and fuzzy logic.

In recent year, neural networks (NNs) [2] became an important technique for classification tasks. Numerous researches done in neural classification demonstrated that neural networks are promising substitutes for many statistical approaches. But the negative side of NNS is that it is hard to understand the importance of neurons and the functions of weights. In [3], fuzzy entropy was used for partitioning of the input feature space into decision regions and for determination of important features with satisfactory separability in the solving of classification

problems. In the literature [4, 5, 6, 7, 8, 9, 10, 11, 12, 13], it is reported that neural fuzzy networks (NFNs) use the good sides of both neural networks and fuzzy systems for classification problems to obtain better performances, different from the pure neural networks or fuzzy systems being used separately [13].

Fuzzy ARTMAP (FAM) [14] and Simplified Fuzzy ARTMAP (SFAM) [15] are two distinctive kinds of neural networks (NNs) that can perform incremental learning. The benefit of these two networks is lower training time compared to other NNs such as Multilayer Perceptron (MLP). These two networks have been widely used in different classification problems [14, 15, 16, 17, 18, 19, 20]. FAM network is composed of three different sections: Fuzzy ARTa, Fuzzy ARTb and Inter ART. Main goal of SFAM networks is classification, what makes it different from FAM, since Fuzzy ARTb section is terminated and not used [17].

SFAM network classification performances are influenced by the presentation order of training patterns. This issue was resolved by training SFAM network couple of times by means of training patterns found in random order. After SFAM is trained several times, the test patterns' expected classes are deposited. The last class prediction for the test patterns is performed by majority voting [14]. Beside the voting results, the average classification results of test patterns from all simulations being performed are also stated [20]. The discussion made in this paper based on SFAM network may possibly be applied to FAM network if it is used for solving classification problems as well.

To prove the efficiency of SFAM network, we selected three different datasets. These are Iris dataset, Wisconsin Breast Cancer dataset and Wine dataset. All these datasets were taken from the University of California at Irvine (UCI) machine learning (ML) database. These datasets are commonly used among in literature where Machine Learning methods were used employed for classification tasks what gave us an ability to compare the performance of our proposed automated classification system with other system prosed in other researched related with these three classification problems.

This paper is organized as follows. A literature review is introduced in Section 2 where SFAM network was explained in more details. Section 3 presents the experimental results and discussion. Section 4 provides the conclusions and possible future improvements.

2. SIMPLIFIED FUZZY ARTMAP NETWORK

SFAM is a fast, online, supervised learning method which uses simple fuzzy rules such as min-max for neuron activation and selection. The advantage of using fuzzy rules is that they minimize computation necessary for learning and they can learn each pattern with very small

number of iterations. SFAM begins having no any connection weight, enlarges in size to fit the task, employs simple learning equations, and contains just parameter selected by user recognized as vigilance parameter [21].

SFAM network is composed of two parts (Fuzzy ART and Inter ART). These two sections build firm recognition categories as the answer to the series of inputs. In the stage of supervised learning, Fuzzy ART takes a stream of input pattern features that relate to the output groups in the category layer. Inter ART section functions by raising the Fuzzy ART's vigilance parameter (VP) slightly to decrease a predictive error in the output category layer. The role of VP is to adjust the lowest confidence that Fuzzy ART must have in an input vector so that Fuzzy ART is capable of taking that category, instead of finding for a better one via an automatically controlled procedure of hypothesis analysis. [14, 20].

Fuzzy ART section is composed of three different layers: L_0 , L_1 and L_2 . These three layers consist of nodes. Nodes in the first layer L_0 form input vector and L_0 activity vector can be expressed as $P = (P_1, \dots, P_M)$. Each of these values must be normalized and in interval $[0, 1]$ [20].

Rise of categories can be evaded in the case that inputs are normalized using the complement coding technique. For that reason, the complement coded input P to the layer L_1 is the $2M$ dimensional vector [20]

$$P = (a, a^c) \text{ where } a_i^c = 1 - a_i \quad (1)$$

Nodes in the L_1 layer are linked to nodes in the output layer F_2 with a weight vector. For all L_2 category nodes j ($j = 1, \dots, N$), there exist a weight vector related to nodes in layer L_1 $w_j = (w_{j1}, \dots, w_{j2M})$ with adaptive weights.

The initial condition is [20]:

$$w_{j1}(0) = \dots = w_{j,2M}(0) = 1 \quad (2)$$

meaning that all groups are indifferent.

For all input patterns P and L_2 node j , the choice function T_j is

$$T_j(P) = \frac{|P \wedge w_j|}{\alpha + |w_j|} \quad (3)$$

where the fuzzy AND operator \wedge is defined as

$$(m \wedge n)_i = \min(m_i, n_i) \quad (4)$$

and the $norm$ is

$$|m| = \sum_{i=1}^M |m_i| \quad (5)$$

for any M -dimensional vectors m and n . To simplify, let, $T_j(I)$ in Equation (3) be represented as T_j if the input I is

fixed. A category choice is done when one F_2 node starts being active at a certain time. The category choice can be indicated by J , where [20]

$$T_j = \max \{T_j : j = 1, \dots, N\} \quad (6)$$

In the case that more than one T_j is maximal, the category having a lesser index is selected [20].

Resonance will happen in the case that the match function, $|P \wedge w_j|/|P|$ of the selected group satisfies the vigilance criterion:

$$\frac{|P \wedge w_j|}{|P|} \geq \rho \quad (7)$$

where ρ stands for VP. When resonance occurs, learning begins, as it is described now. Mismatch reset happens if condition given in (7) is not satisfied. If that is a case, then the value of the choice function T_j is 0 and a new index J is selected by Equation (6) [20].

The search process will be pursued till the selected J meets (7). After this processed is finished, the weight vector w_j is changed based on the equation

$$w_j^{(new)} = (P \wedge w_j^{(old)}) \quad (8)$$

with the assumption that fast learning is employed [20].

The Inter ART section will build the mappings among the L_2 and output layer to accurately learn to forecast classification forms. For all inputs, it produces a dynamic weight link composed of many to one or one to one mapping among the L_2 and output layers. Match tracking evades unclearness in forecasting. If this happens, Fuzzy ART search takes either a different category that accurately predicts the aim or free novel category and the dynamic weight link among the Fuzzy ART and Inter ART sections are restructured. Subsequently, VP is changed to the previous VP value. This procedure is pursued till all training patterns are presented [20].

The testing phase operates in the same standard but in this phase there is no match tracking. This is because the input to Fuzzy ART will produce output values at L_2 . The node at L_2 having biggest value will activate the link in the Inter ART section to create the fitting node in output category layer that will represent the forecasted group (class). The order of input patterns in the training phase influences the creation of the nodes at L_2 layer of Fuzzy ART module, even in the case of a fixed VP [20].

3. EXPERIMENTAL RESULTS AND DISCUSSION

A. Experiment 1- Iris Data

The Iris data set [22] (contains 150 patterns that are distributed equally into three output species, Iris Setosa, Iris Versicolor and Iris Virginica. Each pattern consists of four input features: sepal length, sepal width, petal length,

and petal width. We exploit these patterns to produce both the training data and testing data. In the experiment, we used 66 % split that is 100 instances in total were used for training and 50 instances were used for testing. The data set was normalized to the range [0, 1]. The total accuracy obtained in this case was 96 %.

B. Experiment 2 – Wisconsin Breast Cancer Data

Wisconsin breast cancer dataset is second dataset we used for the evaluation of SFAM network. This data set was supported by Dr. William H Wolberg et al. This data can be found in UCI machine learning database. This dataset contains 699 samples obtained from Fine Needle Aspirates (FNA) of person breast tissue and 16 instances have missing values. Since, there are some missing values, these are removed from data set and other 683 cases are used in our experiment. 444 cases are classified as benign and the remaining 239 cases are classified as malignant. Each record in the database has nine attributes. Nine attributes were found to fluctuate notably among benign and malignant instances [23, 24]. All the data was first normalized to range [0, 1]. Half of the total data was used for training phase and another half as used for testing phase. The experiments were performed on 5 different training-test data sets that were obtained via a random process from the original data. Initial VP was set to be 0.75. The average accuracy obtained was 88%. Although this accuracy is not among the highest in the literature it is still satisfactory high. Training and testing time in this case were very minor compared to other methods done in other researches.

C. Experiment 3 – Wine Data

The wine classification data set contains 178 wines that are brewed in the same region of Italy but derived from three different cultivars. Each pattern consists of 13 continuous features: alcohol, malic acid, ash, alkalinity of ash, magnesium, total phenols, flavonoids, nonflavonoid phenols, proanthocyanins, color intensity, hue, OD280/OD315 of diluted wines and proline. First we normalized all the data to the range of [0 1]. Here we used two different approaches for testing and training. In the first approach, we took half of all the data for training phase and another half for testing. This resulted in accuracy of 95.51%. In the second approach, we took 66 % of all data for training and the remaining data for testing. This approach gave us higher accuracy rate. In this case, overall accuracy rate was 96.55 %. Initial vigilance parameter was set to be 0.7.

D. Discussion

Comparison of these three methods is given in Table 1.

The classification was done for different values of VP ranging from 0.1 to 0.9 (in steps of 0.1) but to save space, only the highest performance results for classification were stated in Table 1. The advantage of SFAM network beside giving good overall accuracy is very small amount of time needed for training and testing phase. In [20], genetic algorithm was used to select the presentation order of

training patterns that improves simplified fuzzy ARTMAP classification performance and accuracies obtained in that study were lower than performances obtained in this study. For iris dataset, overall accuracy obtained in [20] was 93.48 % and for wine data accuracy obtained was 95.42 %.

Table 1 Overall performances obtained in conducted experiment

SFAM	Iris Dataset (%)	WBC Dataset (%)	Wine Dataset (%)
Accuracy	96.00	88.00	96.55

In table 2 and 3, we can see comparison of result achieved in this study and results achieved in previous studies where different classifiers were employed for Iris Dataset and Wine data set respectively. As it can be seen, our proposed SFAM network outperformed many of the other classifier methods used in other studies.

Table 2 Comparison of accuracies with previous researches for Iris data set

Author - Years	Methods	Classification Accuracy (%)
Kao et al, 2008 [25]	KNM-PSO	89.93
Li et al, 2008 [26]	Global feature selectionk-means	95.3
Lee et al, 2008 [27]	Iterative fuzzy clustering algorithm	95.4
Chang et. al, 2010 [28]	k-NN	94.6
X. Liu, Y. Ren, 2010 [29]	FCM clustering	89.3
Our Results	SFAM	96.00

Table 3 Comparison of accuracies with previous researches for Wine data set

Author - Years	Methods	Classification Accuracy (%)
Kao et al, 2008 [25]	KNM-PSO	71.63
Chang et. al, 2010 [28]	Neural Networks	95.66
Chang et. al, 2010 [28]	k-NN	92.48
X. Liu, Y. Ren, 2010 [29]	FCM clustering	94.9
B. Gabrys, A. Bargiela, 2000 [30]	General fuzzy min-max NN	92 – 100
Our Results	SFAM	96.55

4. CONCLUSIONS AND FUTURE IMPROVEMENTS

In this study of classification applications, we proposed a Simplified Fuzzy ARTMAP network (SFAM). This study showed that SFAM results in very high performances during classification tasks. The SFAM model is special kind of neural network where neurons are chose and activated by simple fuzzy learning. Fuzzy set theory and fuzzy learning have a significant role when working with uncertainty while doing decisions making tasks in different classification applications. Employing fuzzy learning enabled us to employ the uncertainty in the automated classifier construction and as a result to raise the reliability of the constructed system output. The SFAM network presented in this study is the synergy of the neural network adaptive capabilities and the fuzzy logic. The structures in SFAM networks can be automatically built and free parameters can be attuned by creating online learning schemes simultaneously. The advantage of SFAM network is higher classification accuracy and training and testing time compared to other types of neural networks.

Finally, three different experiments have been employed in this study. The classification results and statistical measures were employed for evaluation of SFAM network performances. The total classification accuracies of the SFAM model were 96.00% in the case of Iris data set, 88 % in the case of WBC dataset and 96.55% in the case of Wine dataset. The results obtained in this study showed that the proposed SFAM network can be used for solving different classification problems efficiently and accurately. As future work, we will try to apply SFAM for different dataset. Special attention will be given to medical classification where we will apply SFAM network to different dataset such as ECG, EEG and EMG datasets.

REFERENCES

- [1] P. O. Duda and P. E. Hart, Pattern classification and scene analysis, New York: Wiley, 1973.
- [2] R. Setiono and H. Liu, "Neural-network feature selector," *IEEE Transactions on Neural Network*, p. 654–662, May 1997.
- [3] H. M. Lee, C. M. Chen, J. M. Chen and Y. L. You, "An efficient fuzzy classifier with feature selection based on fuzzy entropy," *IEEE Transactions on Systems Man, and Cybernetics Part B: Cybernetics*, vol. 31, p. 426–432, June 2001.
- [4] J. S. Jang, "ANFIS: adaptive-network-based fuzzy inference system," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, p. 665–685, 1993.
- [5] C. F. Juang and C. T. Lin, "An online self-constructing neural fuzzy inference network and its

- applications," *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 1, pp. 12-31, 1998.
- [6] C. T. Lin and C. S. G. Lee, *Neural fuzzy systems: A neuro-fuzzy synergism to intelligent system*, Englewood Cliffs, NJ: Prentice-Hall., 1996.
- [7] C. J. Lin and C. T. Lin, "An ART-based fuzzy adaptive learning control network," *IEEE Transactions on Fuzzy Systems*, vol. 5, no. 4, p. 477-496, 1997.
- [8] W. S. Lin, C. H. Tsai and J. S. Liu, "Robust neuro-fuzzy control of multivariable systems by tuning consequent membership functions," *Fuzzy Sets and Systems*, vol. 124, no. 2, p. 181-195, 2001.
- [9] S. Mitra and Y. Hayashi, "Neuro-fuzzy rule generation: Survey in soft computing framework," *IEEE Transactions on Neural Networks*, vol. 11, no. 3, p. 748-768, 2000.
- [10] F. Sun, Z. Sun, L. Li and H. X. Li, "Neuro-fuzzy adaptive control based on dynamic inversion for robotic manipulators," *Fuzzy Sets and Systems*, vol. 134, no. 1, pp. 117-133, 2003.
- [11] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, vol. SMC 15, no. 1, p. 116-132, 1985.
- [12] L. X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples," *IEEE Transactions on Systems, Man, and Cybernetics Part B: Cybernetics*, vol. 22, no. 6, p. 1414-1427, 1992.
- [13] C.-F. Wu, C.-J. Lin and C.-Y. Lee, "A functional neural fuzzy network for classification applications," *Expert Systems with Applications*, vol. 38, p. 6202-6208, 2011.
- [14] G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds and D. B. Rosen, "Fuzzy ARTMAP: a neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Transactions on Neural Networks*, vol. 3, no. 5, p. 698-713, 1992.
- [15] T. Kasuba, "Simplified fuzzy ARTMAP," *AI Expert*, vol. 8, no. 11, p. 19-25, 1993.
- [16] I. Dagher, M. Georgiopoulos, G. L. Heileman and G. Bebis, "An ordering algorithm for pattern presentation in Fuzzy ARTMAP that tends to improve generalization performance," *IEEE Transactions on Neural Networks*, vol. 10, no. 4, p. 768-778, 1999.
- [17] R. Palaniappan, P. Raveendran, S. Nishida and N. Saiwaki, "A new brain-computer interface design using fuzzy ARTMAP," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 10, no. 3, p. 140-148, 2002.
- [18] P. Raveendran, R. Palaniappan and S. Omatu, "Fuzzy ARTMAP classification of invariant features derived using angle of rotation from a neural network," *Information Sciences: An International Journal*, vol. 130, pp. 67-84, 2000.
- [19] M. Vakil-Baghmisheh and N. Pavesic, "A fast simplified fuzzy ARTMAP network," *Neural Processing Letters*, vol. 17, p. 273-316, 2003.
- [20] R. Palaniappan and C. Eswaran, "Using genetic algorithm to select the presentation order of training patterns that improves simplified fuzzy ARTMAP classification performance," *Applied Soft Computing*, vol. 9, p. 100-106, 2009.
- [21] P. Venkatesan and M. L. Suresh, "Classification of Renal Failure Using Simplified Fuzzy Adaptive Resonance Theory Map," *IJCSNS International Journal of Computer Science and Network Security*, vol. 9, no. 11, pp. 129-134, November 2009.
- [22] R. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, p. 179-188, 1936.
- [23] "UCI Machine Learning Repository: Breast Cancer Wisconsin (Diagnostic) Data Set," 2012. [Online]. Available: [http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Diagnostic\)](http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)). [Accessed 15 March 2012].
- [24] "UCI Machine Learning Repository: Breast Cancer Wisconsin (Original) Data Set," 2012. [Online]. Available: [http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Original\)](http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Original)). [Accessed 16 March 2012].
- [25] Y. T. Kao, E. Zahara and L.-W. Kao, "A hybridized approach to data clustering," *Expert Systems with Application*, vol. 34, p. 1754-1762, 2008.
- [26] Y. Li, M. Dong and J. Hua, "Localized feature selection for clustering," *Pattern Recognition Letters*, vol. 29, pp. 10-18, 2008.
- [27] H. Lee, K. Park and Z. Bien, "Iterative fuzzy clustering algorithm with supervision to construct probabilistic fuzzy rule base from numerical data," *IEEE Transactions on Fuzzy Systems*, vol. 2008, pp. 263-277, 2008.
- [28] P.-C. Chang, C.-Y. Fan and W.-Y. Dzan, "A CBR-based fuzzy decision tree approach for database classification," *Expert Systems with Applications*, vol. 37, p. 214-225, 2010.
- [29] X. Liu and R. Ren, "Novel artificial intelligent techniques via AFS theory: Feature selection, concept categorization and characteristic description," *Applied Soft Computing*, vol. 10, p. 793-805, 2010.
- [30] B. Gabrys and A. Bargiela, "General fuzzy min-max neural network for clustering classification," *IEEE Transactions on Neural Networks*, vol. 11, no. 3, p. 2000, 769-783.