Bagged Nonlinear Hebbian Learning Algorithm for Fuzzy Cognitive Maps Working on Classification Tasks

Elpiniki I. Papageorgiou[1], Panagiotis Oikonomou[2], and Arthi Kannappan[3]

[1] Technological Educational Institute of Lamia, Informatics & Computer Technology Department, 3rd km Old National Road Lamia-Arthens, 35100 Lamia, Greece
epapageorgiou@teilam.gr
http://users.teilam.gr/~epapageorgiou
[2] University of Thessaly, Computer and Communication Engineering Dept, 38221 Volos, Greece
paikonom@uth.gr
[3] RVS College of Computer Applications, Coimbatore-641402, Tamilnadu, India
Sek_art@rediffmail.com

Abstract. Learning of fuzzy cognitive maps (FCMs) is one of the most useful characteristics which have a high impact on modeling and inference capabilities of them. The learning approaches for FCMs are concentrated on learning the connection matrix, based either on expert intervention and/or on the available historical data. Most learning approaches for FCMs are Hebbian-based and evolutionary-based algorithms. A new learning algorithm for FCMs is proposed in this research work, inheriting the main aspects of the bagging approach which is an ensemble based learning approach. The FCM nonlinear Hebbian learning (NHL) algorithm enhanced by the bagging technique is investigated contributing to an approach where the model is trained using NHL algorithm as a base learner classifier. This work is inspired from the neural networks ensembles and it is used to learn the FCM ensembles produced by the NHL exploiting better classification accuracies.

1 Introduction

Fuzzy cognitive mapping is a method for analysing and depicting human perception of a given system. The method produces a conceptual model which is not limited by exact values and measurements, and thus it is well suited to represent relatively unstructured knowledge and causalities expressed in imprecise forms. FCM is a dynamic tool because cause-effect relations and feedback mechanisms are involved [1]. The advantageous modeling features of FCMs, such as simplicity, adaptability and capability of approximating abstractive structures encourage us to use them for complex problems in diverse scientific areas [2][3].

In most cases, FCMs are constructed manually, and, thus, they cannot be applied when dealing with large number of variables. In such cases, their development could be significantly affected by the limited knowledge and skills of the knowledge engineer. Thus, it is essential to use learning algorithms to accomplish this task. The

I. Maglogiannis, V. Plagianakos, and I. Vlahavas (Eds.): SETN 2012, LNAI 7297, pp. 157–164, 2012.

adaptive Hebbian-based learning algorithms, the evolutionary-based such as genetic algorithms and the hybrid approaches composed of Hebbian type and genetic algorithm are the most efficient and widely used methods for training FCMs [4-11].

The aim of this study is to present a new ensemble based learning approach using the bagging technique to enhance the learning capabilities of the FCM-NHL approach [5]. In this research work, we investigated the bagging approach for NHL algorithm to learn FCMs. The NHL algorithm was used as a base learning algorithm in bagging ensemble paradigm. The explored learning approach on bagged FCM-NHL is applied to autism classification to show its functionality. The results were compared with the previous ones produced by NHL algorithm, and showed that the learning approach can further be improved using ensemble FCM approach. Ensemble FCM uses FCM trained with NHL algorithm and incorporates the ideas of ensemble method of bagging. Here, the goal of bagged FCM-NHL construction is to achieve better generalization ability over a single classifier in the case of FCMs. The proposed method uses an autistic data set to train FCMs for constructing ensemble FCMs.

2 Ensemble Learning and Fuzzy Cognitive Map

2.1 Ensemble Learning

Ensemble learning is one of the most promising areas of soft computing, which is used successfully in many real world applications such as text categorization, optical character recognition, face recognition and computer aided medical diagnosis [12][13]. The idea of designing ensembles was originated as an alternative way for improving performance of individual classifiers by exploiting knowledge derived from different sources. Ensemble methods overcome the statistical problem, the computational problem and the representation problem of learning algorithms which output is only a single hypothesis, and thus they overcome the limitation of traditional learning algorithms.

In ensemble learning, an agent takes a number of learning algorithms and combines their output to make a prediction. The algorithms being combined are called base-level algorithms [12]. Base-level algorithms are usually generated from training data by a base learning algorithm which can be decision tree, neural network or other kinds of learning algorithms. Most ensemble methods use a single base learning algorithm to produce homogeneous base-level algorithms, but there are also some methods which use multiple learning algorithms to produce heterogeneous learners.

There are many effective ensemble methods, but the most representative ones are Boosting [14] and Bagging [15]. In bagging, if there are m training examples, the base-level algorithms are trained on sets of m randomly drawn, with replacement, sets of the training examples. In each of these sets, some examples are not chosen, and some are duplicated. On average, each set contains about 63% of the original examples. In boosting, there is a sequence of classifiers in which each classifier uses a weighted set of examples. Those examples that the previous classifiers misclassified are weighted more. Weighting examples can either be incorporated into the base-level algorithms or can affect which examples are chosen as training examples for the future classifiers.

2.2 Main Aspects of Fuzzy Cognitive Maps

An FCM is depicted as a fuzzy causal graph [1], in which nodes represent concepts, whereas directed edges between the concepts denote causal relationships present between them. A given system is defined as a collection of concepts (events, actions, values, goals, etc.) that influence each other through cause-effect relationships, which are quantified and usually normalized to the [-1, 1] interval. Positive values describe promoting effect, whereas negative ones describe inhibiting effect. Other values correspond to different intermediate levels of the causal effect. Figure 1 shows a generic representation of the FCM model.

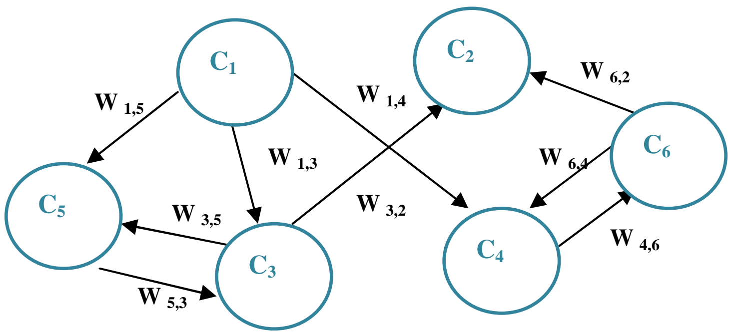

Fig. 1. A Fuzzy Cognitive Map model example

In general, we define FCM as an order pair $< C, W >$, where C is the set of labels and W is the connection matrix. Every label $A_i \in C$ its mapped to its activation value $A_i \in [0,1]$, where 0 means no activation, and 1 means full activation. The labels from C can be interpreted as linguistic terms [1],[2] that point to fuzzy sets. In such case, the activation value A_i is interpreted as the value of fuzzy membership function that measures the degree in which an observed value belongs to the fuzzy set pointed by the related term. The other, simplified interpretation of C can be such that the labels C_i denote the real valued variables, the domains of these variables are assumed as normalized into the [0, 1] interval.

According to the FCM development process, the number and kind of concepts are determined by a group of experts that comprise the FCM model. Then, each interconnection is described by a domain expert either with an if-then rule that infers a fuzzy linguistic variable from a determined set or with a direct fuzzy linguistic weight, which associates the relationship between the two concepts and determines the grade of causality between the two concepts.

In the sequel, all the linguistic variables suggested by experts, are aggregated using the SUM method and an overall linguistic weight is produced. This weight is then transformed into a numerical weight e_{ij}, belonging to the interval [-1, 1], by using the defuzzification method of the center of gravity and finally a numerical weight for e_{ij} is calculated. Using this method, all the weights of the FCM model are inferred.

Once the FCM is constructed, it can receive data from its input concepts, perform reasoning and infer decisions as values of its output concepts. During reasoning the FCM iteratively calculates its state until convergence. The state is represented by a *state vector* A^k, which consists of real node values $A_i(k) \in [0,1]$, i=1,2,…N at an iteration k. The value of each node is calculated by the following equation [16]:

$$A_i(k+1) = f\Big((2A_i(k)-1) + \sum_{\substack{j \neq i \\ j=1}}^{N} (2A_j(k)-1) \cdot E_{ji} \Big) \tag{1}$$

which is a rescaled simulation process that removes the spurious influence of inactive concepts (with C_i = 0) on other concepts, and avoids the conflicts emerge in cases where the initial values of concepts are 0 or 0.5.

2.3 Learning Algorithms for Fuzzy Cognitive Maps

The learning approaches for FCMs are concentrated on learning the connection matrix ***E***, based either on expert intervention and/or on the available historical data. According to the available type of knowledge, the learning techniques can be categorized into three groups: Hebbian-based, population-based and hybrid combining the main aspects of Hebbian-based and evolution-based type learning algorithms. These types of learning algorithms are the most efficient and widely used for training FCMs [4-11]. In the first case, Hebbian-based learning algorithms, such as Active Hebbian Learning (AHL) and NHL were proposed for learning the weight matrix of FCMs based on experts' intervention. On the second case, the experts are substituted by historical data and the underlying learning algorithm used them to estimate the entries of the connection matrix ***E***. In the third case of hybrid learning approaches, the learning goal is to modify/update weight matrices based on initial knowledge from experts and historical data at a two stage process.

A recent review study on learning algorithms for FCMs and their advanced applications is presented in [4]. This study summarizes the main features of the Hebbian type learning, population-based type and hybrid learning for FCMs, depicting at the same time their recent applications in diverse research areas by pinpointing the degree of success of each one.

3 Bagged NHL Learning Approach for Fuzzy Cognitive Map

FCM ensemble is a learning paradigm where a collection of finite number of FCMs is trained for the same task and used for specific application domains. Learning many FCMs and combining their predictions provide an FCM ensemble. The objective of FCM ensemble construction with training different FCMs using NHL, as the base classifier, is to achieve better classification ability over a single FCM classifier. The NHL algorithm has already used to train FCMs for classification tasks showing its functionality [5][169]. The main steps of the NHL algorithm are presented in [16].

The performance of NHL for FCM training can be improved by applying ensemble learning of bagging for solving classification tasks. In this proposed algorithm, since cross validation approach is applied to the data sets, the available data set is partitioned into k subsets with equal size, and then we use part of the k subsets for training while the remaining subset used for performance evaluation. In the examined case study, the subsets are k = 10, thus the data is divided into ten sets, where one set is stored as test data and the remaining used as training data. The value of each concept and the value of weight are updated. Next the test data is trained using the ensemble method and simple majority method is used to calculate the final accuracy. The proposed bagging learning of FCMs along with NHL algorithm, namely bagged NHL-FCM, is described at follows.

Algorithm: Bagging

```
Input:
 1. D is INPUT DATA of 23 concepts given by experts
 2. W is initial weight matrix
 3. Training data SB (total number of records-40)
 4. Weak-learning algorithm for FCMs, NHL
 5. Integer T=5,10, specifying the number of iteration
 6. Integer k=10 specifying the number of subsets (each
    subset has 4 records)
 7. Percent F (70% or 90%) to create resample training
    set St
Do k=1, ..10
 Do t=1, ..T
 8. Take a resample replica St by 70-90% of SB
 9. Call NHL for FCMs with St and receive the classifi-
    er FCM-t which produce after testing an output
    hk(i) for each k test set consisting of i=4 records
 10. Add hk(i) to the ensemble E
 End
End
```

Test: Simple majority voting method –Given a test set k, with $\mathbf{x}_i$, where i=4 records

For i=1, ..4
 For t=1, ..T
  1. Evaluate the ensemble E_i={ht(i)}={h1(i),h2(i), ..hT(i)} for $\mathbf{x}_i$ record
  2. Let $v_{t,j} = \begin{cases} 1, ht - picks - class - \omega_j \\ 0, otherwise \end{cases}$
     be the vote given to class ω_j by classifier ht.
  3. Obtain total vote received by each class
     $V_j = \sum_{t=1}^{T} v_{t,j}$ for j=1,..C in the case of C classes
  4. Choose the class that receives the highest total vote as the final classification.
 End
End

Diversity in bagging is obtained by using bootstrapped replicas of the training data: different training data subsets are randomly drawn—with replacement—from the entire training data [15]. Each training data subset is used to train a different classifier of the same type. Individual classifiers are then combined by taking mean summed output of their decisions (see Figure 2). For any given instance, the class chosen by most classifiers is the ensemble decision.

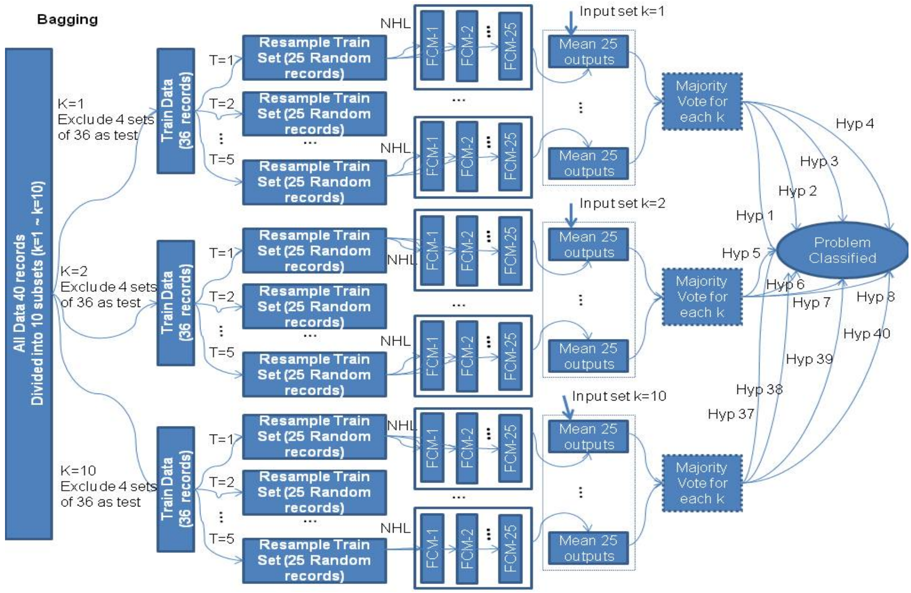

Fig. 2. Bagging approach implemented in FCM-NHL algorithm

Bagging is particularly appealing when available data is of limited size. To ensure that there are sufficient training samples in each subset, relatively large portions of the samples (70% to 90%) are drawn into each subset. This causes individual training subsets to overlap significantly, with many of the same instances appearing in most subsets, and some instances appearing multiple times in a given subset.

4 Experimental Analysis and Results

Autism is a developmental disorder in which attention shifting is known to be restricted. A number of methods have been elicited to study this problem. In a recent work, Arthi K. et al. (2011) analyzed the performance of FCMs on autistic disorder modeling and prediction using NHL algorithm [16]. In that study, the FCM model constructed by physicians to assess three levels of autism (no autism, probable autism and autism) was trained using NHL algorithm for forty real children cases. The produced FCM model consists of 24 concepts. The concept C_{24} has been considered from the experts as DC and could be categorized as Definite Autism (DA), Probable Autism (PA) and No Autism (NA). The overall classification accuracy of that approach was approximately 79% and outperformed other benchmarking machine learning classification techniques [16].

Forty datasets were collected for classification of three different categories, like twenty six as DA, ten as PA and four as NA children. These forty datasets were gathered in [16]. Each one of the three learning approaches was implemented at the 40 records to predict the classification category of each one.

The classification performance results were gathered in Table 1, depicting the confusion matrices of each one learning approach. The classification accuracies of the two examined training approaches for autism disorder assessment was 82.5% and 87.5% (bagging) respectively. The bagged FCM-NHL gave much better results.

Table 1. Confusion matrices & classification accuracies of FCM-NHL and Bagged FCM-NHL

Learning Algorithms	**FCM-NHL**			**Bagged FCM-NHL** (for T=5 and 70% resample)			**Bagged FCM-NHL** (for T=10 and 90% resample)
Confusion matrices	**DA**	**PA**	**NA**	**DA**	**PA**	**NA**	-
DA	24/26	2/26	-	24/26	2/26	-	-
PA	1/10	8/10	1/10	-	7/10	3/10	-
NA	-	3/4	1/4	-	-	4/4	-
True Positive (All)		33/40			35/40		35/40
Accuracy (%)		**82.5%**			**87.5%**		**87.5%**

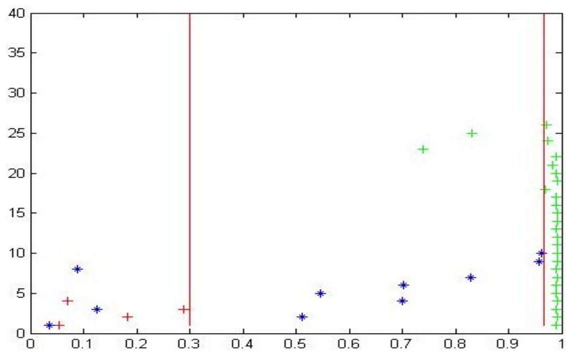

Fig. 3. Classification lines of bagged NHL-FCM algorithm

We also performed a number of experiments for different percentage of resample datasets in the case of bagging, for 70% (25 datasets) and for 90% (33 datasets) and for two different settings of T, for T=5 and T=10. The results were the same for all experimental settings producing correct classes in 35 of 40 cases.

When bagging approaches were implemented, then the 24 records from 26 datasets of "Definite Autism" gave a result of DA and 2 records gave a result of PA. The 7 records from the 10 records of "Probable Autism" gave a result of PA, 0 records were classified as DA and 3 as NA, whereas in the case of "No Autism", the 4 from the 4 records gave the result of NA. Figure 3 presents the classified cases for each one of the three categories. The lines were constructed using the method described in [16]. Applying bagging approach the classification accuracy of FCM-NHL algorithm was improved for the same dataset.

Concluding, the proposed learning methodology was experimentally evaluated in comparison to previous learning algorithm of NHL for autism classification in children exhibiting fast and stable learning. Further research towards a systematic approach to develop FCMs from data could be still carried out. Summarizing, more research work is needed to be done to the extension of the learning capabilities for enhancing and adapting FCM ensembles, as well as the establishment of new FCM learning algorithms.

References

1. Kosko, B.: Fuzzy cognitive maps. Int. J. Man-Machine Studies 24(1), 65–75 (1986)
2. Glykas, G.: Fuzzy Cognitive Maps: Theory, Methodologies, Tools and Applications, 1st edn. Springer, Heidelberg (2010)
3. Papageorgiou, E.I.: A Review Study of FCMs Applications during the last decade. In: Porc. FUZZ-IEEE 2011, Taipei, Taiwan, June 27-30, pp. 828–835 (2011)
4. Papageorgiou, E.I.: Learning Algorithms for Fuzzy Cognitive Maps: A Review Study. IEEE Transactions on SMC Part C (2011) (in press)
5. Papageorgiou, E., Stylios, C., Groumpos, P.: Fuzzy Cognitive Map Learning Based on Nonlinear Hebbian Rule. In: Gedeon, T(T.) D., Fung, L.C.C. (eds.) AI 2003. LNCS (LNAI), vol. 2903, pp. 256–268. Springer, Heidelberg (2003)
6. Froelich, W., Wakulicz-Deja, A.: Mining temporal medical data using adaptive fuzzy cognitive maps. In: Proc. 2nd Conf. on Human System Interactions, HSI 2009, art. no. 5090946, pp. 16–23 (2009)
7. Stach, W., Kurgan, L.A., Pedrycz, W.: M. Reformat, Genetic learning of fuzzy cognitive maps. Fuzzy Sets and Systems 153(3), 371–401 (2005)
8. Papakostas, G.A., Boutalis, Y.S., Koulouriotis, D.E., Mertzios, B.G.: FCMs for pattern recognition applications. Int. J. Pattern Recogn. & Artif. Intel. 22(8), 1461–1486 (2008)
9. Kim, M.-C., Kim, C.O., Hong, S.R., Kwon, I.-H.: Forward-backward analysis of RFID-enabled supply chain using fuzzy cognitive map and genetic algorithm. Expert Systems with Applications 35(3), 1166–1176 (2008)
10. Słoń, G., Yastrebov, A.: Optimization and Adaptation of Dynamic Models of Fuzzy Relational Cognitive Maps. In: Kuznetsov, S.O., Ślęzak, D., Hepting, D.H., Mirkin, B.G. (eds.) RSFDGrC 2011. LNCS, vol. 6743, pp. 95–102. Springer, Heidelberg (2011)
11. Papageorgiou, E.I., Spyridonos, P., Glotsos, D., Stylios, C.D., Ravazoula, P., Nikiforidis, G., Groumpos, P.P.: Brain Tumour Characterization using the Soft. Computing Technique of Fuzzy Cognitive Maps. Applied Soft Computing 8, 820–828 (2008)
12. Policar, R.: Ensemble based systems in decision making, IEEE Circuits and Systems Magazine, third quarter, 21–46 (2006)
13. Zhou, Z.-H.: Ensemble learning. Encyclopedia of Biometrics, 270–273 (2009)
14. Dietterich, T.G.: Machine learning research: Four current directions. AI Magazine 18(4), 97–136 (1997)
15. Breiman, L.: Bagging predictors. Machine Learning 24(2), 123–140 (1996)
16. Kannappan, A., Tamilarasi, A., Papageorgiou, E.I.: Analyzing the performance of fuzzy cognitive maps with non-linear hebbian learning algorithm in predicting autistic disorder. Expert Systems with Applications 38(3), 1282–1292 (2011)