Fuzzy Cognitive Map based Prediction of Pneumonia Severity

Santhi.P and Anto.S

ABSTRACT

Fuzzy Cognitive maps are efficient tools for analysis of decision support systems and time series prediction. In medical domain so many diseases are handled with wrong interventions by the physicians. In this paper FCM based methodology is used for identifying the severity of pneumonia. BigBang-BigCrunch learning method is used for training the FCM. The proposed method is learned with historical data of pneumonia affected patients and tested with real data .The results has less prediction errors compared with conventional genetic algorithms.

Keywords—*Bigbang-BigCrunch optimization, Fuzzy cognitive maps, Learning, prediction*

1 Introduction

Fuzzy Cognitive Maps (FCMs) were originally introduced by Kosko [1] in 1986 as an extension of cognitive maps. They are a convenient modeling tool, usually categorized as a neuro-fuzzy method, for modeling and simulation of dynamic systems. One of their main advantages is an ability to incorporate and adapt human knowledge. FCMs are signed directed graphs used for representing cause-effect relations among nodes or concepts relevant to a system. Concepts are used for describing a system. The behavior of the system is represented by interaction among the nodes. The cause-effect relation among the nodes is determined by expert's knowledge and historical data. They are very flexible and adaptable to a given domain. The domains in which FCMs are applied include electrical engineering, medical science ,supervisory systems,organization and strategy planning,business planning, software project management ,information retrieval,system dynamics and modeling virtual world.

FCM can be constructed manually or by some computational techniques.The manual construction requires atleast one human expert who has good knowledge in the domain.It involves the following steps.

1. Identification of concepts
2. Identification of cause-effect relation among concepts
3. Determine the strength of the relation
4. construct the cognitive map that represent the cause effect relation among the concept with their strength values

P.santhi ,P.G scholar, is with the Department of Computer Science and Engineering, SriKrishna College of Technology,Coimbatore(e-

S.Anto, Assistant Professor, is with the Department of Computer Science and Engineering, SriKrishna College of Technology,Coimbatore (e-mail: santocse@gmail.com).

Sometimes there may be lack of experts or the expert's knowledge may differ from others.FCM designed by each expert may be biased .When the number of nodes increases in a large amount, the FCM development procedure become more complex and difficult. Even if there are available historical data to justify the model's quality, obtaining appropriate model that mimics the data requires laborious effort, which is performed by drawing and simulating successive models.For the above reasons, the development of FCMs in an automated way is necessary.It require learning of FCM.There are three approaches for learning FCMs.ie.,Hebbian based, Global optimization based, and hybrid approaches.

2. Related Works

Previous studies on predicting pneumonia were focused either on rules to classify a patient to a group of risk of getting pneumonia [14] or data mining techniques that extract rules from data to predict pneumonia risk [15]. Artificial neural networks and machine learning techniques were investigated to predict the outcomes of patients with community-acquired pneumonia [15]. However,the previous works that have been done to predict pneumonia state using evolutionary FCMs,[12] which are not black-box methods and are able to deal with fuzziness to describe the relationships among decision variables.

There are few approaches to time series prediction by FCM, learned using different algorithms. The learning goal of FCMs is to compute a weight matrix that best fits the decision-making and prediction problems. Learning algorithms can train FCMs, which means the adjustment of the connections (weights) among concepts. Different learning algorithms for FCMs can be listed as follows

Hebbian based

DHL (Dickerson and Kosko)[6]
AHL (papageorgiou et al)[2]
NHL (papageorgiou et al)[3]
DDNHL (Stach et al)[7]

Population based

Particle Swam Optimization (papageorgiou et al)[4]
Memetic algorithm (Petalas et al)[8]
Genetic algorithm (Mateou et al)[5]
Real Coded genetic algorithm (Stach et al)[11]
Divide and conquer for genetic learning (Stach et al)[10]

One of the main disadvantages of the proposed genetic algorithm based learning methods is its mediocre scalability as the number of parameters to be established grows quadratically with the size of the FCM model. This is because the genetic optimization applied to this modeling is time consuming especially when dealing with large a number of variables. Therefore another approach to speed-up the learning process based on a "divide and conquer" strategy is proposed. The most important problem of the genetic algorithm based learning methods is the selection of the genetic algorithm operator parameters. The convergence and the time of learning depend on this choice. Therefore, there is a need of a new simple global optimization based learning method.

In this study, Big Bang - Big Crunch (BB-BC) optimization method [13]is recommended for learning of FCMs as an alternative to the existing learning methods. This global optimization method is preferred since it has a low computational cost, a high convergence speed and just a few parameters which should be set by the designer for learning of FCMs.

3. Description of medical problem

A disease is usually manifested by symptoms that are expressed by the deviations of observations or medical measurements from their normal state. The proper classification of symptoms leads to the medical diagnosis.On the basis of the diagnosis and the current state of the disease a doctor plans the following treatment. The prescribed treatment can produce unexpected effects that require supplementary interventions that had not been anticipated. The problem of minimizing the risk of wrong interventions made by physicians is addressed in the proposed system.

Pneumonia is an infection of the lower respiratory tract that represents a major cause of mortality worldwide .Laboratory examinations are done to verify the diagnosis of pneumonia.In our case study, in order to proceed to the prediction of medical time series, 10 hospitalized patients with confirmed pneumonia, aged from 19 to 85 years old, most of them men, were selected by our physicians from the hospital's database. For every patient, three measurements per day were accomplished.. Assuming that the current measurements are observed, a physician predicts the possible future state of the patient and on this basis prescribes a medical treatment. i.e.the doctor can make a proper medical intervention, e.g., the prescription of a drug.

Table.1(pneumonia attributes)

attributes	description
X1	temperature
X2	Systolic blood pressure
X3	Diastolic blood pressure
X4	Heartrate
X5	pH
X6	pO2(partial pressure of oxygen)
X7	pCO2(partial pressure of CO2)
X8	HCO3(bicarbonate)
X9	BE(base Excess)
X10	Hb(hemoglobin)
X11	HCT(Hematocrit)
X12	SO2(oxygen saturation)
X13	Na(sodium)
X14	K(potassium)
X15	WBC(white blood cells)

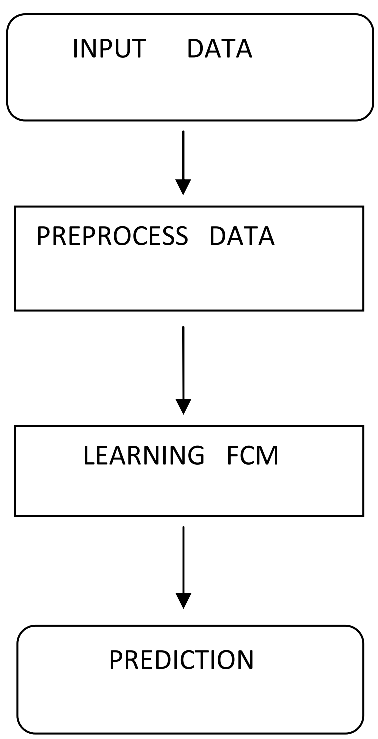
ISSN: 2278-0181
Vol. 1 Issue 8, October - 2012

Pneumonia Sample Investigation table.

	T	SBP	DBP	HR	ph	PO2	PCO2	HCO3	BE	HB	HCt	SO2	TC	Na	K
1	102	90	60	110	7.37	87	38	24	.77	13	40	90	13000	128	3.6
2	100	100	70	104	7.36	90	42	26	.94	14	42	96	5000	136	4.4
3	100	110	70	100	7.35	88	42	23	1.9	12.2	38	98	9800	140	4
4	99.8	130	80	90	7.35	94	43	24	1.0	16	49	99	5400	138	3.9
5	103	90	60	120	7.28	80	46	26	.15	14.6	45.4	88	15600	130	4.7
6	100.5	140	80	96	7.38	91	39	28	2.8	13.2	40.2	99	7800	142	5.1
7	99.5	150	90	84	7.42	90	36	22	1.9	15.8	52	99	6700	136	4.2
8	100	110	70	94	7.36	96	44	22	2.7	10.6	33	97	11200	137	3.8
9	99	130	90	78	7.39	94	34	25	.43	14.4	46	98	4300	141	3.9
10	98.5	130	60	80	7.41	90	36	26	1.6	12.4	41	95	7500	140	4.4

3 Proposed Work

Basic idea in the proposed system can be represented schematically as follows



4. FCM: Mathematical Formulation

Mathematically, a Fuzzy Cognitive Map F is a 4-tuple (C, E, A, f) (Kosko, 1986), where

C = c1,c2,c3………cN where c1 to cN are the N concepts(nodes) forming the nodes of graph

E : (C_i, C_j)---> e_{ij} where e_{ij} is the strength of relation between pair of concepts
A : C_i ---> $C_{i\ (t)}$ where A is the set of activation degrees associated with each concept at time t
recurring relation between C(t+1) and C(t) is written as

$$\forall i \in \{1,\ldots,N\}, C_j(t+1) = f\left(\sum_{i=1}^{N} e_{ij} C_i(t)\right)$$

f is the transformation function which is used to reduce weighted sum within an interval[0,1]

Transformation function used here is logistic

$$f(x) = \frac{1}{1+e^{-Mx}}$$

Where M is the parameter used to determine the degree of fuzzification of function

The construction of FCM as a predictive model requires the selection of attributes *xi* , the optimization of weights *wij*, the definitions of fuzzification $\mu(xi)$, and transformation $f(x)$ functions, respectively

5. BigBang-BigCrunch algorithm

This paper focuses on the BigBang-BigCrunch learning of FCM,which is related with the optimization of the weight matrix *W*.The weights constitute the genotype (the consecutive rows of *W* are placed within the genotype [16]) of the candidate FCM(phenotype). The population of genotypes undergoes the optimization process. The BigBang-BigCrunch optimization algorithms based on the following general procedure . In this algorithm there are two phases; Big Bang phase which is based on Big Bang Theory and Big Crunch phase which is based on Big Crunch Theory. In Big Bang Phase, the candidate's population is randomly generated over the search space and in Big Crunch phase these candidates are concentrated towards center of mass which acts as a convergence operator. Evaluation of new position of each candidate is done using center of mass. This process continues until convergence is obtained [13].In the original version of the algorithm, center of mass is calculated as follows:

$$x_c = \frac{\sum_{i=1}^{N} \frac{1}{f^i} x_i}{\sum_{i=1}^{N} \frac{1}{f^i}}$$

Where
X_{c}= position of center of mass,
X_{i}= position of candidate,
f^{i}= fitness function value of candidate i,
N= population size.
The new generation for next iteration Big Bang phase is normally distributed around x_c. The new candidates around center of mass are calculated by adding or subtracting standard deviation of normal distribution

$$x_i^{new} = x_c + \sigma$$

The standard deviation decreases as the iteration elapse according to following formula

$$\sigma = \frac{\alpha.(x_{max} - x_{min}).rand}{k}$$

Where x_{min}, x_{max} are minimum and maximum range respectively,

rand = random number between 0 and 1,

k = number of iterations.

α = parameter limiting size of search space

Therefore the new candidate is generated as follows:

$X_i^{new} = x_c+\alpha. (x_{max}-x_{min}) .rand/k$

Optimization Algorithm

(Big Bang Phase)

1).Generate randomly N number of candidates with in the search space.

2). Obtain the fitness function values of all candidate solutions.

(Big Crunch Phase)

3). Evaluate the center of mass using equation

4).The new candidates are evaluated around the new point calculated in step 3 by adding or subtracting the standard deviation whose value decreases as the iteration elapse.

5).Return to step 2 until stopping criteria has been met

6. Datapreprocessing

The only fuzzification function that was used till now for learning and testing FCMs was the simple normalization . It is proposed to use the following function for the fuzzification of data as follows:

$$a_i(t) = \frac{x_i(t) - \text{mean}(x_i) + \lambda_i \cdot \text{stdDev}(x_i)}{2\lambda_i \text{stdDev}(x_i)}.$$

The parameter λi determines the length of the interval $[l(xi), u(xi)]$ from which the values of xi are fuzzified. In fact, this goal is accomplished for any parameter λi . However, by decreasing λi , there is a data loss, which makes the learned FCM to be less reliable. After performing multiple experiments, the obtained results showed that due to the data loss, the prediction errors increase for $\lambda i < 4$. For $\lambda i \geq 4$, the prediction errors remain at the same level, and the differences are hardly recognizable . . On the basis of the previous analysis, it was assumed that $\lambda i = 4$.

7. Proposed Learning Methodology

The proposed BB-BC learning method develops a candidate FCM from input data . The input data are given as time series and that consist of a sequence of state vectors which describe a given system at consecutive iteration.The number of these successive iterations of the given historical data is called as the data length. Because of the nature of FCMs, the data points are normalized to the unit interval [0, 1] and they correspond to the degree of presence of a given concept at a particular iteration. Given a system consisting of N concepts, the FCM model can be described fully by its connection matrix. The aim of the learning method is to establish the connection matrix that consists of N(N-1) variables assuming values in [-1, 1]. The proposed method uses the BB-BC algorithm and given input data to determine these values. In other words, the learning goal is to generate the same state vector sequence using the candidate FCM for the same initial vector as it is defined in the input data. Thereby, the candidate FCM generalizes the relations between the concepts, and it allows performing simulations from different initial state vectors in order to represent conclusions about the modeled system. One of the most important considerations for BB-BC,similar to the other global optimization methods, is the design of a cost function, which is appropriate for a given problem. In literature many different cost functions are proposed. In this study the following cost function is found to be appropriate

$$J_1 = \frac{1}{(K-1).N} \sum_{t=1}^{K} \sum_{n=1}^{N} (C_n(t) - \hat{C}_n(t))^2$$

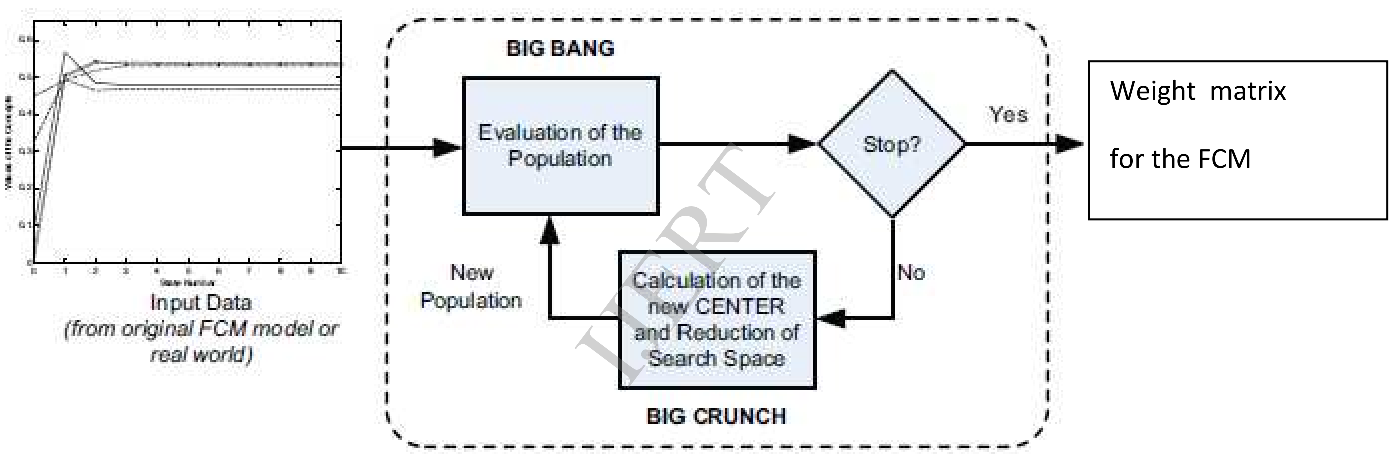
where Cn(t) is the given system response, $\hat{C}_n(t)$ is the candidate FCM response of the n^{th} concept for the initial state vector, K is the data length, and N is the number of concepts.In order to normalize and visualize the cost function the following fitness function which has the value [0 1] is proposed:

$$f = \frac{1}{\theta J_1 + 1}$$

where parameter θ is a positive scaling constant.

Another important condition for the BB-BC learning of FCM is the stopping criteria. This can be defined by following ways:

i. The algorithm terminates if time exceeds the specific time.
ii. The algorithm terminates if the number of BB-BC generations exceeds the specific number.
iii. The algorithm terminates if the best candidate FCM was not improved after a period.
iv. The algorithm terminates if the predefined fitness function value is reached with the best candidate FCM.
In this study, the second stopping criterion (number of generation) is used for experiments.



Learning Algorithm (BB-BC)

7. Results and Discussion

It is possible to claim that the proposed approach really improves the prediction capabilities of FCM. The prediction errors can be significantly decreased which means that the proposed methodology is able to forecast efficiently the state of pneumonia infections and thus to help physicians with medical interventions. The prediction produced by FCM simulates in fact the most likely future state of a patient. The knowledge on the predicted patient state supports a physician in undertaking a medical intervention.The proposed approach would be a submodule of a decision support system in case it is supported by the FCM methodology. Applying the FCMs for time series prediction of real medical data, FCM models that represent and analyze the medical process can be constructed to support the decisions accomplished by physicians.In study [10], where genetic algorithm is proposed as learning method,the system is similar to the ones presented in this paper. Even though a large number of iterations with several populations are needed for GA based learning, in this paper successful results can be obtained just with limited amount of iterations and minimum populations when the proposed BB-BC learning method is used.

ISSN: 2278-0181
Vol. 1 Issue 8, October - 2012

8. Conclusion

In this study, BigBang-BigCrunch based FCM methodology was exploited to address the problem of time-series prediction. The Computational experiments can prove the advantage of using the proposed methodology. The prediction errors calculated by the new FCM approach were lower when compared with the results produced by the conventional genetic-based approaches.In upcoming work, the efforts will focus on further enhancement of the proposed approach investigating methods for learning FCM with some other intelligent techniques like neural network and fuzzy logic.

Acknowledgement

The authors would like to thank the physicians from the medical college of Thrissur ,Dr.RethishBabu, and Dr. Vipin, for their help in data collection

REFERENCES

[1] B. Kosko, "Fuzzy cognitive maps," *Int. J. Man-Mach. Studies*, vol. 24,no. 1, pp. 65–75, 1986.
[2] E. Papageorgiou, C. D. Stylios, and P. P. Groumpos, "Active Hebbian learning algorithm to train fuzzy cognitive maps," *Int. J. Approximate Reasoning*, vol. 37, no. 3, pp. 219–249, 2004.
[3] E. Papageorgiou, C. D. Stylios, and P. P. Groumpos, "Fuzzy cognitive map learning based on nonlinear Hebbian rule," *Australian Conf. Artif. Intell.*, pp. 256–268, 2003.
[4] E. I. Papageorgiou, K. E. Parsopoulos, E. Konstantinos, C. S. Stylios, P. P. Groumpos, and M. N. Vrahatis, "Fuzzy cognitive maps learning using particle swarm optimization," *J. Intell. Inf. Syst.*, vol. 25, pp. 95–121, 2005.
[5] W. Stach, L. A. Kurgan,W. Pedrycz, and M. Reformat, "Genetic learning of fuzzy cognitive maps," *Fuzzy Sets Syst.*, vol. 153, no. 3, pp. 371–401, 2005.
[6] J. A. Dickerson and B. Kosko, "Virtual worlds as fuzzy cognitive maps,"*Presence*, vol. 3, no. 2, pp. 173–189, 1994.
[7] W. Stach, L. Kurgan, and W. Pedrycz, "Data driven nonlinear Hebbian learning method for fuzzy cognitive maps," in *Proc. IEEE World Congr. Comput. Intell.*, Hong Kong, Jun. 1–6, 2008, pp. 1975–1981.
[8] Y. G. Petalas, E. I. Papageorgiou, K. E. Parsopoulos, P. P. Groumpos,and M. N. Vrahatis, "Fuzzy cognitive maps learning using memetic algorithms,"in *Proc. Int. Conf. Comput. Methods Sci. Eng.*, Loutraki, Greece,
2005, pp. 1420–1423.
[9] N. H. Mateou, M. Moiseos, and A. S. Andreou, "Multi-objective evolutionary fuzzy cognitive maps for decision support," in *Proc. IEEE Congr. Evol. Comput.*, vol. 1, Edinburgh, U.K., Sep. 2005, pp. 824–830.
[10] W. Stach, L. Kurgan, and W. Pedrycz, "A divide and conquer method for learning large fuzzy cognitive maps," *Fuzzy Sets Syst.*, vol. 161, no. 19, pp. 2515–2532, Oct. 2010.
[11] W. Stach, L. A. Kurgan,W. Pedrycz, and M. Reformat, "Genetic learning of fuzzy cognitive maps," *Fuzzy Sets Syst.*, vol. 153, no. 3, pp. 371–401,
[12]E. I. Papageorgiou and W. Froelich, "Application of Evolutionary Fuzzy Cognitive Mapsfor Prediction of Pulmonary Infections," in *J.. Inf. Technol.Appl. Biomed.*, January 2012, pp. 1–4.
[13]Big Bang – Big Crunch Learning Method for Fuzzy Cognitive Maps, Engin Yesil, Leon Urbas, World Academy of Science, Engineering and Technology 47 2010
[14] M. J. Fine, T. E. Auble, D. M. Yealy, B. H. Hanusa, L. A. Weissfeld, D. E. Singer, C. M. Coley, T. J. Marrie, and W. N. Kapoor, "A prediction rule to identify low-risk patients with community-acquired pneumonia,"*N. Engl. J. Med.*, vol. 336, no. 4, pp. 243–250, 1997.
[15] P. Heckerling, B. Gerber, T. Tape, and R. Wigton, "Prediction of community acquired pneumonia using artificial neural networks," *Med. Decision Making.*, vol. 23, no. 2, pp. 112–121, 2004.
[16] W. Stach. (2010). Learning and aggregation of fuzzy cognitive maps—An evolutionary approach. Ph.D. Thesis, University of Alberta, Edmonton, AB, Canada. [Online]. Available: http://gradworks. umi.com/NR/62/NR62921.html.