Enhancing the Performance of Neurofuzzy Predictors by Emotional Learning Algorithm

Caro Lucas[1], Ali Abbaspour[2], Ali Gholipour[2], Babak N. Araabi[2], and Mehrdad Fatourechi[3]

[1]Control and Intelligent Processing Center of Excellence, Electrical and Computer Eng. Department, University of Tehran, Tehran, Iran and School of Intelligent Systems, Institute for studies in theoretical Physics and Mathematics, Tehran, Iran lucas@ipm.ir
[2]Control and Intelligent Processing Center of Excellence, Electrical and Computer Eng. Department, University of Tehran, Tehran, Iran
aabbaspr@ut.ac.ir, gholipoor@ut.ac.ir, araabi@ut.ac.ir
[3]Electrical and Computer Eng. Department, University of British Columbia, BC, Canada mehrdadf@ece.ubc.ca

Keywords: Emotional Learning, Prediction, Nonlinear Time Series, Neurofuzzy Model

Abstract

Neural networks and Neurofuzzy models have been successfully used in the prediction of nonlinear time series. Several learning methods have been introduced to train the Neurofuzzy predictors, such as ANFIS, ASMOD and FUREGA. Many of these methods, constructed over Takagi Sugeno fuzzy inference system, are characterized by high generalization. However, they differ in computational complexity. The emotional Learning, which is successfully used in bounded rational decision making, is introduced as an appropriate method to achieve particular goals in the prediction of real world data. For example, predicting the peaks of sunspot numbers (maximum of solar activity) is more important due to its major effects on earth and satellites. The emotional learning based fuzzy inference system (ELFIS) has the advantages of simplicity and low computational complexity in comparison with other multi-objective optimization methods. The efficiency of proposed predictor is shown in two examples of highly nonlinear time series. Appropriate emotional signal is composed for the prediction of solar activity and price of securities. It is observed that ELFIS performs better predictions in the important regions of solar maximum, and is also a fast and efficient algorithm to enhance the performance of ANFIS predictor in both examples.

1 Introduction

Predicting the future has been an interesting important problem in human mind. Alongside great achievements in this endeavor there remain many natural phenomena the successful predictions of which have so far eluded researchers. Some have been proven unpredictable due to the nature of their stochasticity. Others have been shown to be chaotic: with continuous and bounded frequency spectrum resembling white noise and sensitivity to initial conditions attested via positive Lyapunov exponents resulting in long term unpredictability of the time series. There are several developed methods to distinguish chaotic systems from the others, however model-free nonlinear predictors can be used in most cases without changes.

Comparing with the early days of using classical methods like polynomial approximators, neural networks have shown better performance, and even better are their successors: Neurofuzzy models [1], [2], [3], [4]. Some remarkable algorithms have been proposed to train the neurofuzzy models [4], [5], [6], [7]. The pioneers, Takagi and Sugeno, presented an adaptive algorithm for their fuzzy inference system [5]. Some other methods, including adaptive B-spline modeling [6] and adaptive network-based fuzzy inference system [7], fulfill the principle of network parsimony which leads to high generalization of performance. Generalization is the most desired property of a predictor. The principle of parsimony says that the best models are those with the simplest acceptable structures and the smallest number of adjustable parameters.

Following the directions of biologically motivated intelligent computing, the emotional learning methodology has been introduced on the base of emotions which are argued, in contemporary psychology, to be better predictors of future achievements than IQ [8], [9]. The simulated approach is formulated on the base of an emotional signal which shows the emotions of a critic about the overall performance of the system. The emotional signal can be produced by any combination of objectives or goals which improve the estimation or prediction. The loss function will be defined as a function of emotional signal and the training algorithm will be simply designed to minimize this loss function. Thus the need for elaborated definitions of loss function in multi objective problems, which results in high computational complexity, is simply handled by defining an appropriate emotional signal. The cost which should be paid is that the result will be just satisficing rather than optimizing. As a result, the model will be trained to provide the desired performance in a holistic manner. The emotional learning algorithm has three distinctive properties in comparison with other learning methodologies. For one thing, one can use very complicated definitions for emotional signal without increasing the computational complexity of algorithm or worrying about differentiability or renderability into recursive formulation problems. For another, the parameters can be adjusted in a simple intuitive way to obtain the best performance. Besides, the training is very fast and efficient. As can be seen these properties make the method preferable in real time applications like control and decision making, as have been presented in literature [10],[11],[12],[13],[14],[15],[16],[17],[18].

In this research the emotional learning algorithm has been used in the purposeful prediction of some real world data: the sunspot numbers and the price of securities. In predicting the sunspot number time series, the peak points, related to solar maximum regions, are more important to be predicted than the others due to their strong effects on space weather, communication systems and satellites. Additional achievements are fast training of model and low computational complexity. The main contribution of this paper is to provide accurate predictions using emotional learning for Takagi Sugeno neurofuzzy model. The results are compared with other methods of training neural and neurofuzzy models like RBF and ANFIS. The paper consists of six parts; the main aspects of Takagi-Sugeno fuzzy inference system along with associated learning methods are described in the second section. The third section deals with the various forms of utilizing emotional learning in the prediction problem. The results of applying the proposed prediction method to benchmark time series are reported and analyzed in sections four and five. Finally, the last section presents some remarkable properties of emotional learning and some concluding remarks.

2 NeuroFuzzy models

Two major approaches of trainable neurofuzzy models can be distinguished. The network based Takagi-Sugeno fuzzy inference system and the locally linear neurofuzzy model. The locally linear model is equivalent to Takagi-Sugeno fuzzy inference system under certain conditions, and can be interpreted as an extension of normalized RBF network as well [2]. Therefore, the mathematical description of Takagi Sugeno neurofuzzy model which is the most general formulation will be described in this section.

The Takagi-Sugeno fuzzy inference system is constructed by fuzzy rules of the following type

$$\begin{aligned} &Rule_i: \ If\ u_1 = A_{i1}\ And\ \ldots\ And\ u_p = A_{ip} \\ &then\ \hat{y} = f_i(u_1, u_2, \ldots, u_p) \end{aligned} \tag{1}$$

Where $i = 1...M$ and M is the number of fuzzy rules. $u_1,...,u_p$ are the inputs of network, each A_{ij} denotes the fuzzy set for input u_j in rule *i* and $f_i(.)$ is a crisp function which is defined as a linear combination of inputs in most applications

$$\hat{y} = \omega_{i0} + \omega_{i1}u_1 + \omega_{i2}u_2 + \ldots + \omega_{ip}u_p \tag{2}$$

Matrix form $\hat{y} = a^T(\underline{u}) \cdot W$

Thus the output of this model can be calculated by

$$\hat{y} = \frac{\sum_{i=1}^{M} f_i(\underline{u})\mu_i(\underline{u})}{\sum_{i=1}^{M} \mu_i(\underline{u})} \quad ; \quad \mu_i(\underline{u}) = \prod_{j=1}^{p} \mu_{ij}(u_j) \tag{3}$$

Where $\mu_{ij}(u_j)$ is the membership function of *jth* input in the *ith* rule and $\mu_i(\underline{u})$ is the degree of validity of the *ith* rule. This system can be formulated in the basis function realization which clarifies the relation between Takagi-Sugeno fuzzy inference system and the normalized RBF network. The basis function will be

$$\phi_i(\underline{u}) = \frac{\mu_i(\underline{u})}{\sum_{j=1}^{M} \mu_j(\underline{u})} \tag{4}$$

as a result

$$\sum_{j=1}^{M} \phi_j(\underline{u}) = 1 \tag{5}$$

This neurofuzzy model has two sets of adjustable parameters; first the antecedent parameters, which belong to the input membership functions such as centers and deviations of Gaussians; second the rule consequent parameters such as the linear weights of output in equation (2). It is more common to optimize only the rule consequent parameters. This can be simply done by linear techniques like least squares [2]. A linguistic interpretation to determine the antecedent parameters is usually adequate. However, one can opt to use a more powerful nonlinear method to optimize all parameters together. Gradient based learning algorithms can be used in the optimization of consequent linear parameters. Supervised learning is aimed to minimize the following loss function (mean square error of estimation):

$$J = \frac{1}{N} \sum_{i=1}^{N} (y(i) - \hat{y}(i))^2 \tag{6}$$

where *N* is the number of data samples.

According to the matrix form of (2) this loss function can be expanded in the quadratic form

$$J = W^T RW - 2W^T P + Y^T Y / N \tag{7}$$

Where $R = (1/N)A^T A$ is the autocorrelation matrix, *A* is the $N \times p$ solution matrix whose *ith* row is $a(\underline{u}(i))$ and $P = (1/N)A^T y$ is the p dimensional cross correlation vector. From

$$\frac{\partial J}{\partial W} = 2RW - 2P = 0 \tag{8}$$

the following linear equations are obtained to minimize J:

$$RW = P \tag{9}$$

and W is simply defined by pseudo inverse calculation. One of the simplest local nonlinear optimization techniques is the steepest descent. In this method the direction of changes in parameters will be opposite to the gradient of cost function

$$\Delta W(i) = -\frac{\partial J}{\partial W(i)} = 2P - 2RW(i) \tag{10}$$

and

$$W(i+1) = W(i) + \eta \cdot \Delta W(i) \tag{11}$$

where η is the learning rate.

Other nonlinear local optimization techniques can be used for this purpose, e.g. the conjugate gradient or Levenberg-Marquardt which are faster than steepest descent. All these methods have the possibility of getting stuck at local minima. Some of the advanced learning algorithms, that have been proposed for the optimization of parameters in Takagi-Sugeno fuzzy inference system, include ASMOD (Adaptive B-Spline modeling of observation data) [6], ANFIS (Adaptive network based fuzzy inference system) [7] and FUREGA (fuzzy rule extraction by genetic algorithm) [2]. ANFIS is one of the most popular algorithms that has been used for different purposes, such as system identification, control, prediction and signal processing. It is a hybrid learning method based on gradient descent and least square estimation. ASMOD is an additive constructive algorithm based on *k*-d tree partitioning. It reduces the problems of derivative computation, because of the favorable properties of B-spline basis functions. Although ASMOD has a complicated procedure, it has advantages like high generalization and accurate estimation.

One of the most important problems in learning is the prevention of over fitting. It can be done by observing the error index of test data at learning iterations. The learning algorithm will be terminated, when the error index of test data starts to increase, in an average sense. Prevention of over fitting is the most common way of providing high generalization.

3 Emotional Learning

Satisficing approaches to decision making has, is recent years, been widely adopted for dealing with complex engineering problems [18]. New learning algorithms like reinforcement learning, Q-learning, and the method of temporal differences [19], [20], [21], [22], [23] are characterized by their fast computation and in some cases lower error in comparison with classical learning methods. They can be interpreted as approximations to dynamic programming, which although furnishes a well known computational algorithm, via recursive solution of the Bellman-Jacobean-Hamilton equation and perhaps the best example of fully rational approach to decision making, is notorious for its computational complexity, sometimes referred to as the “curse of dimensionality” [24], [25]. Fast training is a notable consideration in control applications. Prediction applications also belong to the class of decision making problems where two desired characteristics are accuracy and low computational complexity.

The Emotional learning method is a psychologically motivated algorithm which is developed to reduce the complexity of computations in prediction problems with particular goals. In this method the reinforcement signal is replaced by an emotional cue, which can be interpreted as a cognitive assessment of the present state in light of goals and intentions. The main reason of using emotion in a prediction problem is to lower the prediction error in some regions or according to some features. For example predicting the sunspot number is more important in the peak points of the eleven-year cycle of solar activity, or accurate prediction of the peaks and valleys in the price of securities may be desired. This method is based on an emotional signal which shows the emotions of a critic about the overall performance of prediction. The emotional signal can be produced by any combination of objectives or goals which improve estimation or prediction. The loss function will be defined just as a function of emotional signal and the training algorithm will be simply designed to decrease this loss function. So the predictor will be trained to provide the desired performance in a holistic manner. If the critic emphasizes on some regions or some properties, this can be observed in his emotions and simply affects the characteristics of predictor. Thus the definition of emotional signal is absolutely problem dependent. It can be a function of error, rate of error change and many other features. Finding an appropriate formulation for emotion is not usually possible; in contrast a linguistic fuzzy definition of it is absolutely intuitive and plausible.

A loss function is defined on the base of emotional signal. A simple form is

$$J = \frac{1}{2} K \sum_{i=1}^{N} es(i)^2 \tag{12}$$

where *es(i)* is the of emotional signal to the *ith* sample of training data, and K is a weighting matrix, which can be simply replaced by unity.

Learning is adjusting the weights of model by means of a nonlinear optimization method, e.g. the steepest descent or conjugate gradient. With steepest descent, the weights are adjusted by the following variations:

$$\Delta\omega = -\eta \frac{\partial J}{\partial \omega} \tag{13}$$

where η is the learning rate of the corresponding neurofuzzy controller and the right hand side can be calculated by chain rule:

$$\frac{\partial J}{\partial \omega} = \frac{\partial J}{\partial es} \cdot \frac{\partial es}{\partial y} \cdot \frac{\partial y}{\partial \omega} \tag{14}$$

According to (12): $\frac{\partial J}{\partial es} = K.es$

and $\frac{\partial y}{\partial \omega}$ is accessible from (3) where $f_i(.)$ is a linear function of weights.

Calculating the remaining part, $\frac{\partial es}{\partial y}$, is not straightforward in most cases. This is the price to be paid for the freedom to choose any desired emotional cue as well as not having to impose presuppose any predefined model. However, it can be approximated via simplifying assumptions. If, for example error is defined by

$$e = y_r - y \tag{15}$$

where y_r is the output to be estimated, then

$$\frac{\partial es}{\partial y} = -\frac{\partial es}{\partial e} \tag{16}$$

can be replaced by its sign (-1) in (14). The algorithm is after all, supposed to be satisficing rather than optimizing.

Finally the weights will be updated by the following formula:

$$\Delta\omega = -K \cdot \eta \cdot es \cdot \frac{\partial y}{\partial \omega} = -K \cdot \eta \cdot es \cdot \frac{\sum_{i=1}^{M} u_i \mu_i(\underline{u})}{\sum_{i=1}^{M} \mu_i(\underline{u})} \tag{17}$$

The definition of emotional signal and the gradient based optimization of the emotional learning algorithm in neurofuzzy predictors are clarified among two examples in the next sections.

4 Predicting the Sunspot numbers

Solar activity has major effects not only on satellites and space missions but also on communications and weather on earth. This activity level changes with a period of eleven years, called solar cycle. The solar cycle consists of an active part, the solar maximum, and a quiet part, solar minimum. During the solar maximum there are many sunspots, solar flares and coronal mass ejections. A useful measure of solar activity is the observed sunspot numbers. Sunspots are dark spots on the surface of the sun which last for several days. The SESC sunspot number is computed according to the Wolf's sunspot number $R=k(10g+s)$, where g is the number of sunspot groups, s is the total number of spots in all the groups and k is a variable scaling factor that indicates the conditions of observation.

A variety of techniques have been used in the prediction of solar activity, most of which are based on the sunspot number time series. The sunspot number, which has been saved since 1700, shows low dimensional chaotic behavior and its prediction has been a challenging problem for researchers. However, good results are obtained by methods proposed in several articles [26], [27], [28], [29], [30]. In this research, both the monthly and the yearly averaged sunspot numbers are used to be predicted. Figure 1 shows the history of solar cycles on the base of yearly sunspot numbers. The error index in predicting sunspot numbers, similar to most of the previous studies, is the normalized mean square error (NMSE):

$$NMSE = \left(\frac{\sum_{i=1}^{n} (y - \hat{y})^2}{\sum_{i=1}^{n} (y - \bar{y})^2} \right) \tag{19}$$

In which $y, \hat{y}$ *and* $\bar{y}$ are observed data, predicted data and the average of observed data respectively.

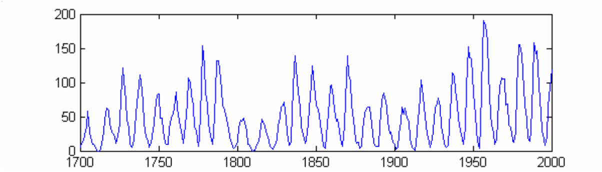

Figure 1: The yearly averaged sunspot number

As the first observation, the emotional learning algorithm has been used to enhance the performance of a neurofuzzy predictor, initially trained by ANFIS. The emotional signal is computed by a linguistic fuzzy inference system with error and rate of error change as inputs. Five and three Gaussian membership functions, negative large, negative, zero, positive and positive large, are used for the inputs (error and rate of error change, respectively) and the emotional signal is calculated by a center of average defuzzifier from the rule base depicted by the surface in figure 2.

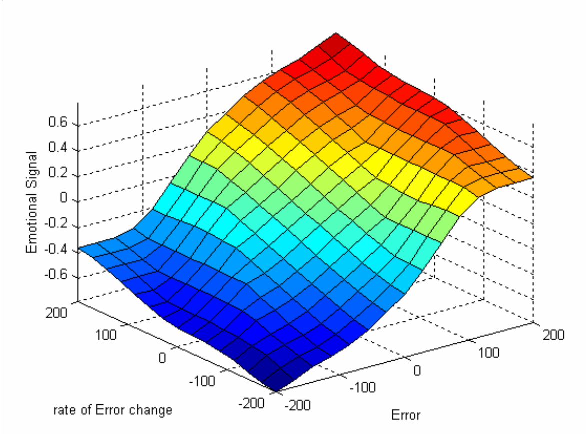

Figure 2: The surface generated by linguistic fuzzy rules of the emotional critic

There are seven Gaussian membership functions for the emotional signal as the output of fuzzy critic. The simulated fuzzy definition of the critic is motivated from our knowledge of emotions in human, and can be extended by inserting more inputs to the system. Figure 3 presents the targeted and predicted outputs of the test set (from 1920 to 2000). The lower diagram shows the results of best fitted data by ANFIS. The training is done with optimal number of fuzzy rules and epochs (74 epochs) and has been continued until the error of validation set had been started to increase.

The other diagram shows the targeted and predicted values after using emotional learning. The emotional algorithm is used in one pass of the training data to fine tune the weights of neurofuzzy model which has been initially adjusted by ANFIS. The error index, NMSE, has been decreased from 0.1429 to 0.0853 after using emotional learning. The improvement of prediction accuracy, especially among the solar maximum regions, is noticeable. It's interesting that training ANFIS to the optimum performance takes approximately ten times more computation effort than the emotional learning to improve the prediction. Thus combining ANFIS with the emotional learning is a fast efficient method to improve the quality of predictions, at least in this example.

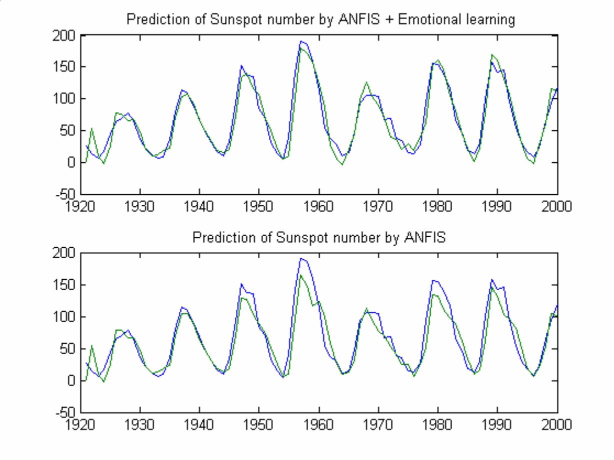

Figure 3: Enhancement in the prediction of sunspot numbers by emotional learning, applied to ANFIS: Targeted and predicted values; lower: by ANFIS, upper: by ANFIS + Emotional Learning

The next results are reported as a comparison of the quality of predicting the monthly sunspot numbers by the Emotional Learning based Fuzzy Inference System (ELFIS) with some other learning methods, the orthogonal least squares learning for the RBF network and Adaptive Network based Fuzzy Inference System (ANFIS). All methods are used in their optimal performance. Over fitting is prevented by observing the mean square error of several validation sets during training. ELFIS is constructed over Takagi Sugeno fuzzy inference system. The emotional signal is computed by a fuzzy critic whose linguistic rules are defined by means of error, rate of error change and the last targeted output. By defining appropriate membership functions for each of the inputs and 45 linguistic fuzzy rules, the desired behavior of emotional critic is provided to show exaggerated emotions in the solar maximum regions. Figure 4 shows the surface generated by the fuzzy rules among the two dimensional space of the more important inputs (prediction error and last observed value of sunspot number). The emotional signal is used as the input to the learning formula (17) where the weights of neurofuzzy model (2) are adjusted. Just three Sugeno type fuzzy rules, like (1), are used in ELFIS to comply with the principle of parsimony. As a result, the matrix of adjustable weights has 9 elements (three weights for the three inputs of each rule). The specifications of methods, NMSE of predictions and computation times (on a 533 MHz Celeron processor) are presented in Table 1. It is observed that learning in ELFIS is at least four times faster than the others and is more accurate than ANFIS. Note that using a functional description of emotional signal rather than the fuzzy description generates a faster algorithm, but finding such a suitable function is not easy.

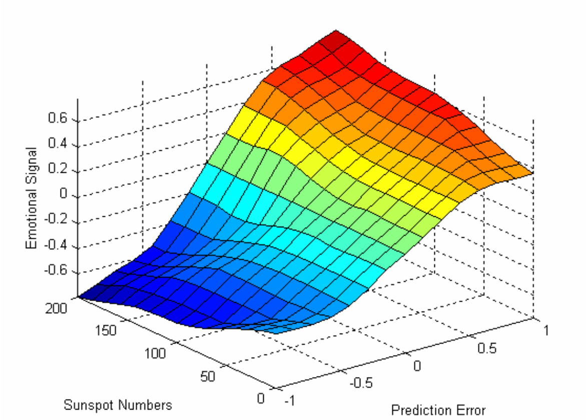

Figure 4: The surface generated by linguistic fuzzy rules of the emotional critic in ELFIS; in the prediction of monthly sunspot numbers

Table 1: Comparison of predictions by selected neural and neurofuzzy models

	Specifications	Computation Time	NMSE
ANFIS	8 rules and 165 epochs	89.5790 sec.	0.1702
RBF	7 neurons in hidden layer	84.7820 sec.	0.1314
ELFIS	3 Sugeno type fuzzy rules	22.3320 sec.	0.1386

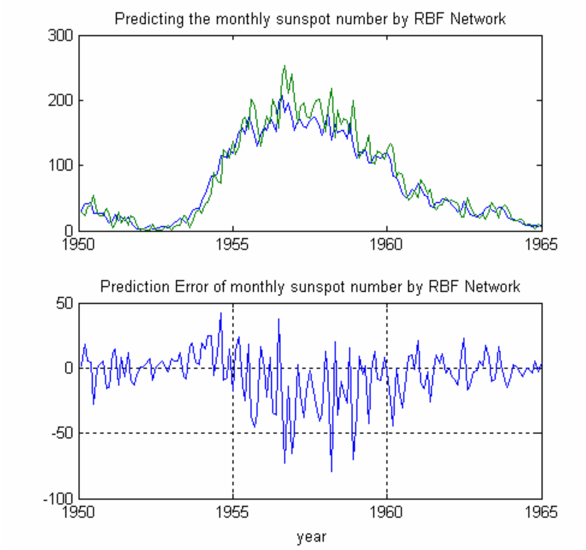

Figure 5: Predicting the sunspot numbers by RBF

Figures 5 to 7 show the predictions by RBF network, ANFIS and ELFIS respectively. These diagrams are a part of test set, especially the cycle 19 which has an above average peak in 1957. It's observable that ELFIS generates the most accurate predictions in the maximum region; however, the NMSE of RBF is the least, indicating that RBF generates more accurate predictions through the total test set. By modifying the validation sets affecting the stop time of learning procedure, even better NMSEs can be obtained in RBF, but this results in higher prediction errors especially in 1957.

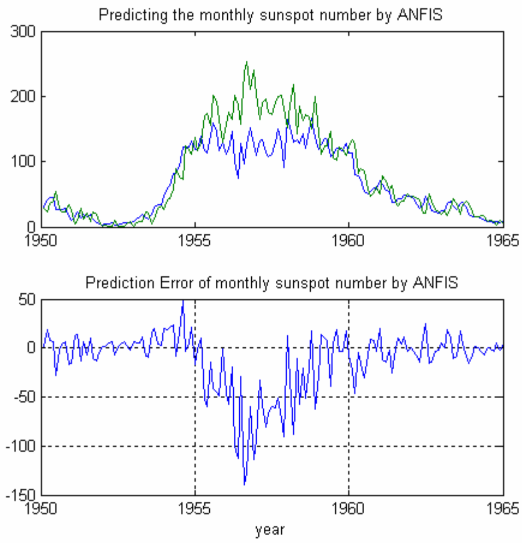

Figure 6: Predicting the monthly sunspot numbers by ANFIS

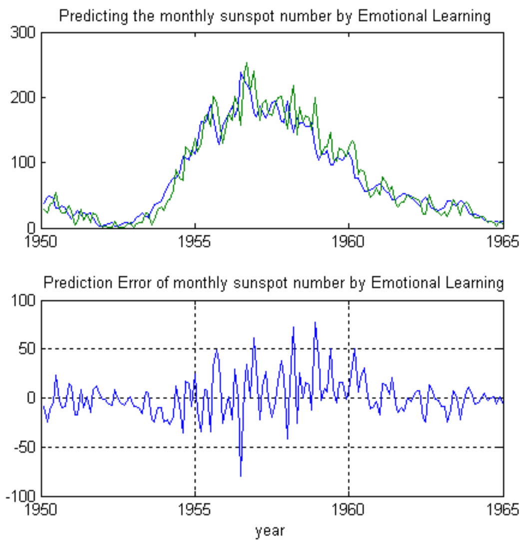

Figure 7: Predicting the monthly sunspot numbers by ELFIS

5 Predicting the Security Price

The second example is the prediction of securities such as stocks, treasury bonds and government bonds; etc. If there is a predictor that predicts the future exactly; then the best investment is on the maximum rate of return. For this reason, the performance of prediction is significant. Some researchers have used neural networks e.g. MLP and RBF for the prediction of securities. In this research, the emotional learning algorithm is applied to the network initially trained by ANFIS to predict the stock price of General Electric (GE) in S&P index 500. For this case one can use various definitions of emotional signal, as a function of prediction error and differential of error, or even any significant event like crossing of spot price with some well monitored moving average. Here the emotional signal is taken as the output of a linguistic fuzzy inference system with the error and the rate of error change as inputs. Five and three Gaussian membership functions are used for the inputs respectively. Figure 8 shows the surface generated by the fuzzy rules of emotional critic.

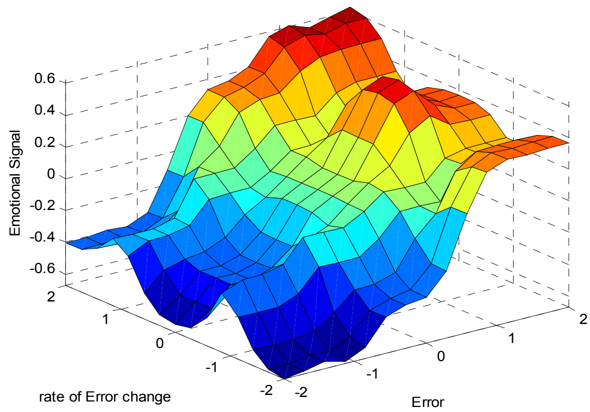

Figure 8: The output surface of a linguistic fuzzy inference system for producing emotional signal.

In this research, the daily closed price for the stock is considered. The model parameters, number of regressors and number of neurons are optimized to prevent over fitting. The stock price of 800 days and the price of 400 following days are used for train data and test data respectively. The result of predicting the stock price by ELFIS is presented in figure 9.

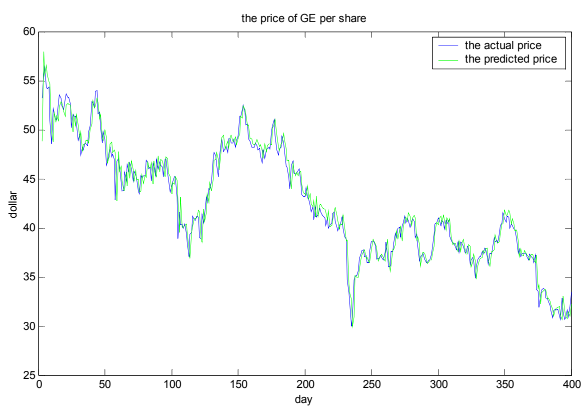

Figure 9: Predicting the security price using emotional learning plus ANFIS

Table 2 presents a comparison of the quality of predicting the daily closed stock price of General Electric (GE) by ELFIS with some other networks such as ANFIS, RBF, and MLP. These results are obtained by considering the over fitting and the optimal neurons in the hidden layer (on a 1.8 GHz Celeron processor). As this practical example shows, the emotional learning algorithm provides more accurate predictions with lower computational complexity.

Table 2: A comparison of various neural and neurofuzzy models in the stock price prediction

	Specifications	Computation Time	NMSE
MLP	37 neurons in hidden layer	6.5600 sec.	0.0347
ANFIS	12 rules and 257 epochs	13.8390 sec.	0.0370
RBF	31 neurons in hidden layer	7.2000 sec.	0.0395
ELFIS	12 Sugeno type fuzzy rules	1.8320 sec.	0.0358

6 Conclusion

Training a system to make decision in the presence of uncertainties is a difficult problem especially when computational resources are limited. Supervised training can not be used because the desired values for the decision variables are unknown. However, the desirability of past decisions can usually be assessed after the outcomes of their implementations are observed. Therefore, unsupervised training methods that do not utilize those assessments can not take full advantage of the available knowledge. Several approximate methods like back propagation through plant, and identification of the plant or (pseudo) inverse plant model have been successfully used in the past couple of decades [31], [32], [33], [34]. Behavioral and emotional approaches to control and decision making can also be classified in this category [35]. Besides providing biological plausibility, they have the extra advantage of not being confined to cheap control problems like set point tracking [10]. The emotional approach is a step higher in the cognitive ladder and can be more useful in goal-aware or context-aware applications (e.g. dealing with multiple objectives in decision problems even when the objectives are fuzzy or can not be differentiated or directly evaluated with simple mathematical expressions).

The main contribution of this paper is the application of those ideas to prediction domain. Although prediction is easier to deal with because we do not have the further complexity of unknown plant, and so the proposed learning methods should also be compared to error minimization methodologies, model free prediction has also become of great importance in the past few decades, and there have been many past efforts to train neuro- and/or fuzzy predictors with alternative loss functions. In this paper, we have used the emotional learning interpolation to two very important benchmark problems. The motivation is not confined to achieving computational efficiency or improving the total prediction accuracy. In both problems, achieving more accurate results in desired regions or according to some important features is an important goal towards which some increase in error indices alongside the total test set can be tolerated. Specifically, one wishes to improve the prediction quality of solar activity (the sunspot number time series) in solar maximum regions (the peak points of sunspot number) at the expense of the prediction accuracy in less interesting regions. In the case of stock market prediction too, the quality of predictions in trend reversal regions (peaks and valleys) are of greater importance for supporting investment decisions.

The achievements reported in this paper are twofold. On the one hand, excellent prediction quality has been achieved for the two different benchmark problems with considerable reduction in computational complexity. On the other hand, a psychologically motivated framework for considering alternative or even multiple goals in decision making (in this case prediction) has been proposed, which is easy to apply even when the goal can not be expressed via well known mathematical expression, or is not differentiable. The goals are satisfied by tuning the predictor so that an emotional signal indicating how the present state is assessed to be non-conductive to the goals, is continually minimized (i.e. we shift gradually to states assessed as more satisfactory with respect to the goals). The proposed emotional learning based fuzzy inference system (ELFIS) has been used in the prediction of solar activity (the sunspot number time series) where the emotional signal is determined with emphasis on the solar maximum regions (the peak points of sunspot number) and it has shown better results in comparison with RBF network and ANFIS. In the prediction of security price, the emotional learning algorithm is defined by emotions of a fuzzy critic and results in good predictions. In fact the use of a combination of error and rate of error change leads to late overtraining of the model. The definition of emotional signal is an important problem in emotional learning algorithm and provides higher degrees of freedom. In the prediction of security price, better performance can be obtained through the use of variables in addition to the lagged values of the process to be predicted (e.g. fundamentalist as well as chartist data).

References:

[1] Brown M., Harris C.J. (1994), *Neuro fuzzy adaptive modeling and control*, Prentice Hall.

[2] Nelles O. (2001), *Nonlinear system identification*, Springer Verlag, Berlin.

[3] Bossley K.M. (1997) *Neurofuzzy Modeling Approaches in System Identification*, PhD thesis, University of Southampton, Southampton, UK
[4] Leung H., Lo T., Wang S. (2001) Prediction of noisy chaotic time series using an optimal radial basis function neural network, *IEEE Tran. On Neural Networks*, 12(5), pp. 1163-1172.
[5] Takagi T., Sugeno M. (1985), Fuzzy identification of systems and its applications to modeling and control, *IEEE Tran. On systems, Man and Cybernetics*, vol. 15, pp. 116-132.
[6]. Kavli T. (1993) ASMOD: An algorithm for adaptive spline modeling of observation data, *Int. J. of Control*, 58(4), pp. 947-967.
[7] Jang J.R. (1993) ANFIS: Adaptive network based fuzzy inference system, *IEEE Tran. On systems, Man and Cybernetics*, 23(3), pp. 665-685.
[8] Goleman D. (1995) *Emotional Intelligence*, New York: Bantam Books.
[9] Picard R.W., Vyzas E., Healey J. (2001) Toward machine emotional intelligence: Analysis of affective physiological state, *IEEE Trans. on Pattern Analysis and Machine Intelligence,* 23 (10), pp. 1175-1191.
[10] Lucas C., Shahmirzadi D., Sheikholeslami N. (2003) Introducing BELBIC: Brain Emotional Learning Based Intelligent Controller, *Accepted for publication in the International Journal of Intelligent Automation and Soft Computing (Autosoft).*
[11] Jazbi A., Lucas C. (1999) Intelligent control with emotional learning, *7th Iranian Conference on Electrical Engineering, ICEE'99,* Tehran, Iran, pp. 207-212.
[12] Lucas C., Jazbi S.A., Fatourechi M., Farshad M. (2000) Cognitive Action Selection with Neurocontrollers, *Third Irano-Armenian Workshop on Neural Networks*, Yerevan, Armenia.
[13] Fatourechi M., Lucas C., Khaki Sedigh A. (2001) An Agent-based Approach to Multivariable Control, *Proc. of IASTED International Conference on Artificial Intelligence and Applications,* Marbella, Spain, pp. 376-381.
[14] Fatourechi M., Lucas C., Khaki Sedigh A. (2001) Reducing Control Effort by means of Emotional Learning, *Proc. of 9th Iranian Conf. on Electrical Engineering ICEE'01,* Tehran, Iran, pp. 41-1 to 41-8.
[15] Fatourechi M., Lucas C., Khaki Sedigh A. (2001) Reduction of Maximum Overshoot by means of Emotional Learning, *Proceedings of 6th Annual CSI Computer Conference*, Isfahan, Iran, pp. 460-467.
[16] Perlovsky L.I. (1999), Emotions, Learning and control, *proc. of IEEE Int. symp. On Intelligent control/Intelligent systems and semiotics,* Cambridge MA, pp. 132-137.
[17] Ventura R., Pinto Ferreira C. (1999) Emotion based control systems, *proc. of IEEE Int. symp. On Intelligent control/Intelligent systems and semiotics,* Cambridge MA, pp. 64-66.
[18] Inoue K., Kawabata K., Kobayashi H. (1996) On a Decision Making System with Emotion, *proc. 5th IEEE International Workshop on Robot and Human Communication*, pp. 461-465.
[19] Sutton R.S., Barto A.G. (1998) *Introduction to reinforcement learning*, MIT press, Cambridge.
[20] Barto A., Sutton R., Watkins C. (1990) Learning and sequential decision making, in *Learning and Computational Neuroscience*, MIT press, Cambridge.
[21] Watkins C. (1989) *Learning from delayed rewards*, PhD thesis, University of Cambridge, UK.
[22] Watkins C., Dayan P. (1992) Q-Learning, *Machine Learning*, *8*, pp. 279-292.
[23] Sutton R.S. (1989) Learning to predict by the method of temporal differences, *Machine Learning*, *3*, pp. 9-44.
[24] Ungar L.H. (2002) Reinforcement learning from limited observations, *Workshop on Learning and Approximate Dynamic Programming,* Playacar, Mexico.
[25] Brown M., Bossley K.M., Mills D.J., Harris C.J. (1995) High dimensional neurofuzzy systems: Overcoming the curse of dimensionality, *proc. of IEEE International Conference on Fuzzy Systems*, New Orleans, USA, pp. 2139-2146.
[26] Izeman A. J. (1985) Wolf J.R. and the Zurich sunspot relative numbers, *The Mathematical Intelligence*, vol.7, no.1, pp. 27-33.
[27] Tong H., Lim K. (1980) Threshold Autoregressive limit cycles and cyclical data, *J. Roy. Statistics. Soc. B*, no.42, pp. 245-292.
[28] Tong H. (1996) *Nonlinear time series: A dynamical system approach*, Oxford press, UK.
[29] Weigend A., Huberman B., Rumelhart D.E. (1990) Predicting the future: a connectionist approach, *Int. J. Of Neural systems*, vol. 1, pp. 193-209.
[30] Weigend A., Huberman B., Rumelhart D.E., (1992) Predicting sunspots and exchange rates with connectionist networks, in *Nonlinear Modeling and Forecasting*, Casdagli, Eubank: Editors, Addison-Wesley, pp. 395-432.
[31] Werbos P.J. (1974) *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*, PhD thesis, Harvard University, USA.
[32] Rumelhart D.E., Hinton G.E., Williams R.J. (1986) Learning internal representations by error propagation, in D.E. Rumelhart and J.L. McClelland, editors, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1, chapter 8, MIT Press, Cambridge.
[33] Narendra K.S., Parthasarathy K. (1990) Identification and Control of dynamical systems using neural networks, *IEEE Tran. On Neural Networks*, 1(1), pp. 4-27.
[34] Xu X., He H.G., Hu D. (2002) Efficient Reinforcement Learning Using Recursive Least-Squares Methods, *Journal of Artificial Intelligence Research*, 16, pp. 259-292.
[35] Bay J.S. (1997) Behavior Learning in Large Homogeneous Populations of Robots, *IASTED International Conference on Artificial Intelligence and Soft Computing,* pp. 137-140.