THE USE OF ANALYTIC HIERARCHY PROCESS FOR PREDICTION

İbrahim Erşan
Department of Computer Engineering
Girne American University
TURKEY.
ibrahimersan@gau.edu.tr

Ali Haydar
Department of Computer Engineering
Girne American University
TURKEY.
ahaydar@gau.edu.tr

ABSTRACT

In literature Analytic Hierarchy Process (AHP) is used for decision making. Reliability of the AHP in decision making is proven with many applications. In many prediction processes, Back-Propagation Multi-Layer Perceptron (BPMLP) is a proven and reliable method. A new system for prediction is proposed which uses basic principles of AHP with alterations on criteria and alternative weight determination. Comparison of the prediction results obtained from well known BPMLP and modified AHP showed that; it is possible to use AHP framework for prediction as well.

Keywords:AHP, ANN, BPMLP, Decision Making, Prediction

INTRODUCTION

Published materials show that Analytic Hierarchy Process (AHP) has a very wide range of usage in almost all branches of real life. Because of its flexible structure and easily applicable mathematical formulation, this process is extensively used in the literature especially in last two decades (Schmoldt et al., 2001)(Rabbani&Rabbani, 1996). AHP is a structured approach for solution of complex decision making problems (Saaty, 1980). It is developed in 1970's by Thomas L. Saaty. Basic principle of AHP depends on ranking priorities of criteria affecting choice of an alternative amongst several alternatives. All criteria and sub-criteria are located in a hierarchy tree to obtain a structured solution.

It is possible to find a lot of applications in literature, for classification, prediction, function-approximation, and image recognition etc., carried out with Artificial Neural Networks (ANN). Especially Back-Propagation Multi-Layer Perceptron (BPMLP) is a widely used method in prediction processes (Haykin, 1999).

In this study, we applied the AHP for prediction and its performance is compared with the results obtained by BPMLP. Same set of data are used for both methods and similar performances are obtained.

CONTEXT AND REVIEW OF LITERATURE

Analytic Hierarchy Process

Any decision making operation which has multiple quantifiable criteria can be carried out with AHP if it is possible to locate these criteria in a hierarchy. This definition seems that AHP covers almost all decision making processes (Saaty, 1994) (Saaty, 2008).

In AHP, decision making is decomposed into four steps, (i) Problem is defined and all information domains that will be used for decision making is determined. Definition of problem is used to determine goal of entire process and information domains are used as criteria. (ii) A hierarchy tree is created by locating goal to the topmost level (Figure 1). Criteria take place in the following level(s) and it is possible for each criterion to have sub-criteria as much as required. In last level alternatives takes place. In figure 1, there are j alternatives and one of these alternatives will be selected as goal, as result of entire process.

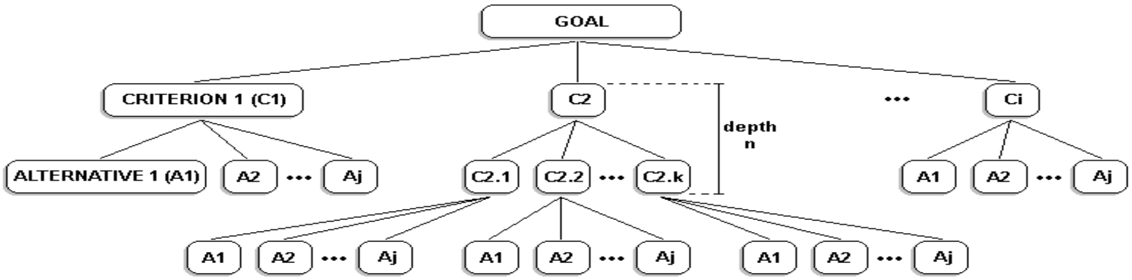

Figure 1.AHP Hierarchy Tree

(iii) A set of matrices is constructed to perform comparison of criterion in hierarchy tree. While creating these matrices a pair-wise approach is considered and each criterion is compared one by one with another. For i criterion, $(i^2-i)/2$ comparison is carried out and results are located in an (i x i) matrix. (iv) While doing pair-wise comparisons, a priority table is used to determine weights (Table 1) (Saaty, June 2008).

Table 1.Priority Table for Pair-Wise Comparison

Intensity Of Importance	Definition	Explanation
1	Equal Importance	Two elements contribute equally to the objective
3	Moderate Imp.	Experience and judgment slightly favor one element over another
5	Strong Importance	Experience and judgment strongly favor one element over another
7	Very Strong Imp.	An element is favored very strongly over another; its dominance demonstrated in practice
9	Extreme Imp.	The evidence favoring one activity over another is of the highest possible order of affirmation
Intensities of 2,4,6 and 8 can be used to express intermediate values. Intensities 1.1, 1.2, 1.3, etc. can be used for elements that are very close in importance.		

Back Propagation Multi-Layer Perceptron

The BPMLP algorithm is a type of supervised, error correction learning that calculates an error on the output layer and propagates that error backwards through the network to determine how each individual weight factor contributes to the output error. In this paper, we used a BPMLP algorithm with a single hidden layer that consists of N1 neurons. The steepest-descent gradient approach used by the BPMLP to minimize the mean square error function (Cichoclar and Unbehaven, 1993) is defined as:

$$E_p = \frac{1}{2}\sum_{j=1}^{n}\left(d_{jp} - y_{jp}\right)^2 = \frac{1}{2}\sum_{j=1}^{n} e_{jp}^2 \quad (1)$$

where, $d_{jp}$is the desired output signal of the *j*th output neuron for the *p*th example, *n* is the number of output neurons and $y_{jp}$is the actual output signal.

The total error function is defined as:

$$E_T = \sum_p E_p = \frac{1}{2}\sum_p \sum_j \left(d_{jp} - y_{jp}\right)^2 \quad (2)$$

For each learning example, the synaptic weights w_{ij} are changed by an amount of

$$\Delta w_{ij} = -\eta \frac{\partial E_p}{\partial w_{ij}}, \quad \eta > 0 \quad (3)$$

The general formula for updating the weights is

$$\Delta w_{ij} = \eta \delta_j o_i \tag{4}$$

where η is the learning rate , ∂_j is the local gradient of the hidden neuron j and o_i is the function signal at the output neuron i.

The data from the input neurons is propagated through the network via the interconnections such that every neuron in a layer is connected to every neuron in the adjacent layers. Each interconnection has associated with it a scalar weight, which acts to modify the strength of the signal passing through it. The neurons within the hidden layer perform two tasks: they sum the weighted inputs to the neuron and then pass the resulting summation through a nonlinear activation function.

The unipolar sigmoid activation function with its output in the range (0, 1) used in this study is as follows:

$$y_j = \varphi(u_j) = \frac{1}{1 + \exp(-\gamma_j u_j)} \tag{5}$$

where, $\varphi(.)$ is the unipolar sigmoid activation function and u_j is defined as the weighted sum of inputs together with its bias value θ_j and is obtained using the formula

$$u_j = \sum_{i=1}^{n} w_{ji} x_i + \theta_j \tag{6}$$

In the above equation a bias is included in order to shift the space of the nonlinearity.

The three parameters, namely the learning parameter η, the sigmoid constant γ, and the number of neurons in hidden layer 1, N1 should be selected in such a way that the total error is minimized.

METHOD

In conventional assessment approach, academic success is evaluated by considering grades taken from exams. In this study, a questionnaire form is prepared with total 64 questions to evaluate academic success of university students without asking any question about grades taken from any exam. In prepared questionnaire, questions are categorized as (i) General academic notion, (ii) Social life, (iii) Unmanageable factors and, (iv) Economical factors. Academic success of students is evaluated with respect to determined categories. Result of conventional assessment, which is known as CGPA, is divided into 5 ranges (0.00-1.39, 1.40-1.69, 1.70-1.99, 2.00-2.49, 2.50-4.00) and taken as base measurement criterion. All questionnaires are divided into two with ratio 1:2. 34% of all questionnaires are separated for testing and 66% for training of system.

Question categories are located to AHP hierarchy tree as criteria and CGPA ranges are located as alternatives. Goal is determined as 'Academic Performance in terms of CGPA'. In order to find correct alternative for a student, priorities of determined categories must be chosen correctly. The one who can determine these priorities most correctly can only be the student himself. By using answers of students in questionnaires it is possible to rank categories with correct priorities. By this way for each student same hierarchy tree with different priority rank can be found which can give correct alternative. Answers of questions in each category are taken as a vector. Distances between these vectors are found and proportions of these distances are used to find priorities of categories. Explained method leads to a prediction system rather than a decision system. In fact prediction is a result of decision.

Obtained results must be compared with results of a reliable method in order to be sure that new method can be used for prediction. For this purpose, ANN is also used for the same data. In this study, the number of inputs to BPMLP is 64 (number of questions in prepared questionnaire) and the number of outputs are 5 (CGPA ranges).

NUMERICAL RESULTS

ANN is a non-linear method and iterations are started with random initial weights so after each learning process different success percentages can be obtained for testing and training data. Best obtained testing data success percentage with ANN is 57.8%.

AHP is a linear method and there is no randomness in any phase, so there is no need to do several iterations. Obtained testing data success percentage is 57.8% and training data success percentage is 82.8%. Following table (Table 2) shows several learning process results of ANN that can be compared with results of AHP. In each learning process some parameters of network is modified to obtain better results.

Table 2.Results obtained using BPMLP

	Trial 1	Trial 2	Trial 3
Momentum Updating	0.9	0.9	0.9
Sigmoid Constant	4	1	2
Number of Epochs	64	39	77
Success on Training Data	62.5%	75.8%	92.2%
Success on Testing Data	57.8%	57.8%	57.8%

DISCUSSION AND CONCLUSION

Main criterion for success of method can be accepted as success rate on testing data. Proposed prediction method with AHP provides same success rate on testing data compared to BPMLP. The experimental results have shown that AHP can be a future promising method for prediction. Furthermore, in AHP, the linear comparison of the vectors can be replaced by a non-linear comparison which may lead better results. For the data at hand, the results obtained are giving exactly the same performance as BPMLP for the testing data and a comparable result for the training data.

REFERENCES

Saaty, T.L., (1980). *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. Newyork: McGraw-Hill, Inc.

Cichoclar, A., Unbehaven, R., (1993), *Neural Networks for Optimization and Signal Processing*, John Wiley & Sons.

Haykin**,** S.**,** (1999), *Neural Networks, Second Edition*, Prentice Hall.

Saaty, T.L., (June 2008). *Relative Measurement and its Generalization in Decision Making: Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors - The Analytic Hierarchy/Network Process*. RACSAM (Review of the Royal Spanish Academy of Sciences, Series A, Mathematics),102(2), p.251–318.

Saaty, T.L., (2008). *Decision Making with the Analytic Hierarchy Process*.Int. J. Services Sciences, 1(1), p.83-98.

Saaty, T.L., (1994) *'How to make a decision: the analytic hierarchy process'*, Interfaces, 24(6), p.19–43.

Schmoldt, D., Kangas, J., Mendoza, G., Pesonen, M., (2001) *The Analytic Hierarchy Process in Natural Resource and Environmental Decision Making*, Managing Forest Ecosystems, Dordrecht, Boston, London: Kluwer Academic Publishers., 3

Rabbani, S.J.R., Rabbani, S.R., (1996) *Decisions in Transportation with the Analytic Hierarchy Process*, Campina Grande, Brazil: Federal University of Paraiba.