BUILDING DESIGN SUPPORT BY HIERARCHICAL EXPERT NETWORKS

Özer Ciftcioglu, Sanja Durmisevic and Sevil Sariyildiz
TU Delft, Faculty of Architecture, Department of Computer Science

ABSTRACT: For building design, computational intelligence systems use a knowledge base formed by means of neural network and fuzzy logic (neuro-fuzzy) techniques, from a building design database. The application of such a system to a building design task was preliminarily demonstrated earlier. The present research describes a systematic neural fuzzy modelling of data that forms a knowledge base in a hierarchical form. Each sub-knowledge base represents a local expert i.e., level-one expert, and the association of local experts forms a more comprehensive expert, i.e., a global domain expert. The association of the experts is accomplished by means of fuzzy-logic-driven gating network that performs the information handling as required. Although, the present paper describes hierarchical experts as local and global, the associations can be done in more subtle form, i.e., in a more than two steps so that the level of experts can be categorized in multi-level form. In such more complex structures, multi-level experts require a related gating network that could similarly be designed.

The building design support system with the expert network developed, as a whole, is generic enough for decision-makings with a novel systematic approach concept using an appropriate database. Accordingly, the research deals with a particular architectural building design with efficiency and consistency features using the hierarchical expert network system described.

KEYWORDS: *Building design, hierarchical experts*, *neural networks, fuzzy logic*

1. INTRODUCTION

Building design is a complex process and it is one of the major activities in architecture. Complexity is due to various components of the process, which should be cared for in a consistent and constructive way throughout the duration from the very beginning up to the realization phase. The components include the persons in charge for the realization as well as the people/bodies concerned with respect to their due involvement and the materials used during the construction. In such a case, the information demand and correspondingly information made available can be even overwhelming and therefore appropriate information processing tools must be used for dealing with this information. This situation is getting increasingly tight due to the technological advances in the Information and Communication Technology (ICT). The means for information acquisition is becoming more effective and distribution of the processed and/or unprocessed information is becoming easier. However, as a result of this, next to demand of fast processing, optimality with respect to routing of the information flow to the prospective users and optimality with respect to the depth of information related to the user of this information are essential considerations for cost effectiveness. Building design involves multi-dimensional aspects to be considered with conflicting criteria. As result of this, many types of expertise are required. It also requires flexibility to accommodate the probable emerging demands in the course of the execution of the building design project. Added to complexity, the latter feature makes the building design also a dynamic process. In the terminology of architecture, the systematic distribution of graded information is termed as information ordering. Here the graded information is the particular information in the form of subsets so that the end user receives not the total information but a sufficient part of it in the sense that the end user can make use of it while redundancy is minimized. Hence next to the task of information processing, information conveyance with without redundancy is important issue in dealing with building design.

Rapid advances in parallel processing technologies have given essential impetus to intelligent information processing, which has become the driving source of an emerging technology known as soft computing. This calls for intelligent systems that are able to process information which may be complex, uncertain even incomplete or contradictory. In this context, neural networks and fuzzy logic are the essential tools. Considering the merits of each approach separately, the most suitable computational intelligence method can be selected for a specific application. Such methods are rather promising in the building sector since the building design information may be complex, uncertain, even incomplete or contradictory. These are all due to the dynamic nature of the design process. The paper describes the use of soft computing technology in building design where soft computing methods are used in a modular form to be able to handle a large volume of data. The organization of the paper is as follows. Part 2 gives a brief description of fuzzy logic and neural networks as equivalent to fuzzy inference systems. Part 3 describes mixture of experts (ME) networks. Part 4 describes ME network as a fuzzy inference system for building design support and presents an application to building data with performance evaluation. Part 5 gives some discussion followed by conclusions.

2. MEANS OF SOFT COMPUTING

2.1 . Fuzzy Logic: An overview

As the building design is a highly knowledge intensive problem, most of the modern building design problems are either too complex or too ill defined to analyze with conventional methods. However, by defining the technical and functional requirements as a fuzzy set, one can perform inexact reasoning during the conceptual or creative phase of the design process with optimal information routing and design decisions. The brief description of fuzzy logic is given below.

Fuzzy set theory and fuzzy inference systems have been introduced by Zadeh (1965,1973). Fuzzy logic explicitly aims to model the imprecise form of human reasoning and decision making. These are essential to our ability to make rational decisions in situations of uncertainty. We encounter such imprecise cases often we encounter in real life situations. We encounter Human reasoning can utilise imprecise propositions, and also infer imprecise consequences. An example of such reasoning can be exemplified by a car driving using the form "***if*** the speed is high ***then*** reduce gas". This heuristic rule does not specify at exactly what point the speed becomes high, nor does it specify the amount by which the speed is reduced. Yet it is still possible to apply this rule to satisfactorily control the speed of the car.

The fundamental concept of fuzzy logic is known as *linguistic variable*. A linguistic variable is a variable, that takes values from spoken language. Considering the above example of driving a car, such a variable can be assigned as high, low, or medium. Although these values do not have precise meaning, a certain distribution between zero and one can be defined and associated with the values. Thus, a speed of 40 km/h can be defined as medium assigning the value 1 for this speed. Any speed around the medium speed of 40 km/h is also medium but the degree of being medium will vary and will be less than that assigned to 40 km/h.The more the speed differs from 40 km/h in any direction, the less the degree of association will be. Such a distribution is commonly referred to as the *membership function* of the linguistic variables. These linguistic variables are called fuzzy variables.

The universe of discourse of a fuzzy variable is the finite input space for which the membership functions are defined. The shape of the membership functions is dependent on the attributes of the underlying concept, and can be represented by any normalized function. Each point in the input space has a degree of membership, which defines the degree to which that point belongs to a given fuzzy value. The membership value is conventionally shown by μ=[0., 1.]. Figure 1 represents the distributions of four linguistic values of speed using trapezoidal functions as fuzzy sets. These are the fuzzy membership functions and the universe of discourse is [0.0, 80] for this particular example.

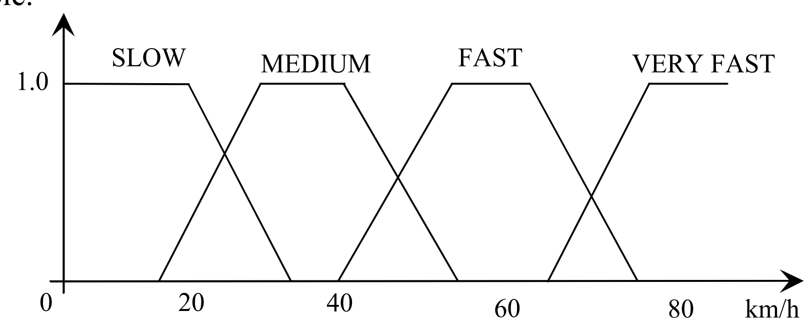

Figure 1: Typical fuzzy sets of speed

The concept of approximate reasoning plays an essential role in fuzzy systems. Typically, fuzzy reasoning is specified by a generalised modus ponens:

if a=A then b=B;
given a=A;
what is b?

All the values in the expressions above are represented by fuzzy membership functions and the implication b is derived using the fuzzy rule termed as *compositional rule of inference.* Conceptually, fuzzy systems are implicitly or explicitly rule-based systems, which comprise rules of the form:

$$
\begin{aligned}
&IF \;\; a_1 = A_1^1 \quad AND \quad a_2 = A_2^1 \;\; AND \; ..... \;\; THEN \;\; b = B_1 \\
&ALSO \\
&IF \;\; a_1 = A_1^2 \quad AND \quad a_2 = A_2^2 \;\; AND \; ..... \;\; THEN \;\; b = B_2 \\
&ALSO \; ....
\end{aligned}
$$

where all variables and values are fuzzy.
One sees from figure 1, that with fuzzy sets, a numerical value is classified into one or more linguistic labels. These labels may be discrete or continuous and they are the membership

functions that represent the numerical strength of linguistic labels for the domain of classification. Since the membership functions can overlap, this results in multi-value representation of the knowledge. An input value intersects with one or more membership functions of the input classification and therefore it is attached to several linguistic labels.

2.2. Radial Basis Functions Network: An overview

A basic radial basis function (RBF) network may be depicted as shown in Figure 2. Without loss of generality, the number of outputs in the network can be extended to a multi-output case. The architecture consists of an input layer, a hidden layer and an output layer. The hidden layer consists of a set of radial basis functions as nodes. Each node has a parameter vector **c** defining a cluster center whose dimension is equal to the input vector. The hidden layer node calculates the Euclidean distance between the center and the network's input vector. The calculated distance calculated is used to determine the radial base function output. Conventionally, all the radial basis functions in the hidden layer nodes are of the same type and usually gaussian. The response of the output layer node(s) can be seen as a map f: $R^n \rightarrow R$, of the form

$$f(\mathbf{x}) = \Sigma\, \mathbf{w}_i\, \Phi(||\mathbf{x} - \mathbf{c}_i||)^2 \qquad (1)$$

Here the summation is over the number of training data N. $\mathbf{c}_i$ (i=1,2,….,N) is the i-th center which may be equal to the input vector $\mathbf{x}_i$ or may be determined in some other way. $\mathbf{w}_i$ is the weight vector of the i-th center.

Once the basis function outputs are determined, the connection weights from the hidden layer to the output are determined from a linear set of equations. As a result, accurate functional approximation is obtained. Complexity increases as the size of the training data increases. For a large data set this may become unpractical. Therefore it is desirable to use a limited number of hidden layer nodes in place of having a number equal to N. In the present description, for sake of simplicity in representation and description, a single function is considered so that neural network has one output for each multivariable input. For this case the output is given by

$$f(x) = w_o + \sum_{j=1}^{N} w_i \phi(\| x - c_j \|) \qquad (2)$$

where $x \in R^p$ is the input vector.

The functional equivalence of radial basis function networks and fuzzy inference systems is already established (Jang and Sun 1993; Hunt et al. 1996). In this equivalence the normalized gaussian functions play the role of membership functions as shown in part 4. The RBF network has rather appealing properties for soft computing. Next to their equivalence to fuzzy inference systems under lenient conditions, they can be used for multivariable functional approximation using basis functions, and they can also be considered as a type of artificial neural network.

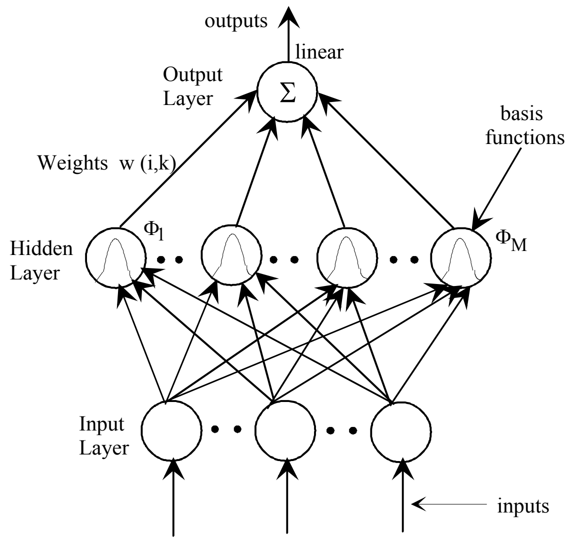

Figure 2: A basic RBF network architecture

3. MIXTURE OF EXPERTS NETWORK

3.1. Modular networks

In the terminology of neural networks, a network is said to be modular if the computation performed by the network can be decomposed into two or more modules, each of which working using distinct inputs without communicating with each other. A modular neural network architecture devised in this way is called Mixture of Experts (ME) (Jacobs et al., 1991). It consists of several expert networks trained on different partitions of the input space. The output of expert networks are combined by a gating network simultaneously trained in order to weight the experts' output according to their performance at solving the problem. Here, each network is a local expert, i.e., level-one expert, and the associations of local experts forms a more comprehensive expert that becomes a level-two global domain expert as level-two. Hierarchically, the associations can be done in a more subtle form, i.e., in more than two steps so that the level of experts can be categorised in a multi-level form. In such more complex structures, multi-level experts require a related gating network that could similarly be designed. A two-level mixture of experts network and its expansion for multi-level representation is shown in figure 3 (a) and (b), respectively.

Conventionally the mixture of experts is considered from the probabilistic viewpoint. Accordingly, a special type of training algorithm called Expectation-Maximization (EM) have been introduced by Jordan and Jacobs (1994) and its convergence properties are investigated by Jordan and Xu (1995). In EM approach, the gating network is simultaneously trained in order to select stochastically the expert that performs the best for a given task.

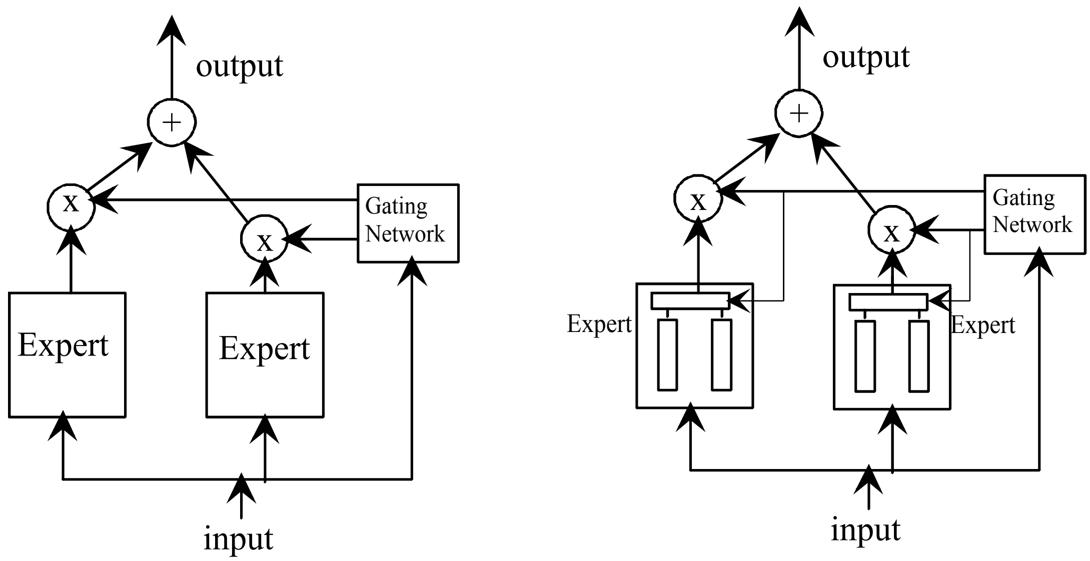

Figure 3 : Mixture of experts network and its multi-level form

3.2. Fuzzy Mixture of Experts Networks

Considering building design, the task is carried out by soft computing methods such as the fuzzy logic. This approach has appealing features for building design information. In the present work, the gating network is a fuzzy inference system in contrast with a probabilistic network as shown in figure 4. The fuzzy gating network, in principle, operates on IF-Then rules a fuzzy rule base can be expressed by

If **x** is R_i Then y is Expert Model_i

The gating network correspond to the "**if**"" part of the If-Then rule . The R_i node is the rule node that computes the matching degree of the input x for rule i. It can be implemented as a soft weighting factor that may be a gaussian-like function

$$\mu_{pj}(x_p) = \exp(-(x_p - x_{op})^2 / 2\sigma_{pj}^2 \qquad (3)$$

where x_{op} and σ_{pj} are the mean and variance of the gaussian, respectively. A fuzzy "And" is performed by arithmetic multiplication. After a normalisation process a weighting factor for the outputs of the local experts is obtained which is the membership function μ_{pj} of the fuzzy sets. The output fuzzy set O_i from the i-th expert model is a crisp value, which is a fuzzy singleton. The weighting factor having been determined by the if part, the defuzzification is computed by taking the weighted contributions from different model outputs. The output y is determined by the centroid defuzzification method as

$$y = \sum_{i=1}^{i=n} o_i q_i(x) \qquad (4)$$

where

$$q_i = \frac{\mu_i}{\sum_{p=1}^{n} \mu_p} \tag{5}$$

and

$$\mu_p = \prod_{j=1}^{n} \mu_{jp}(x_j) \tag{6}$$

In figure 4, the nodes after the rule nodes R_1,.., R_m normalize the matching degrees from the rule nodes to produce the weighting factor for the outputs of the local experts.

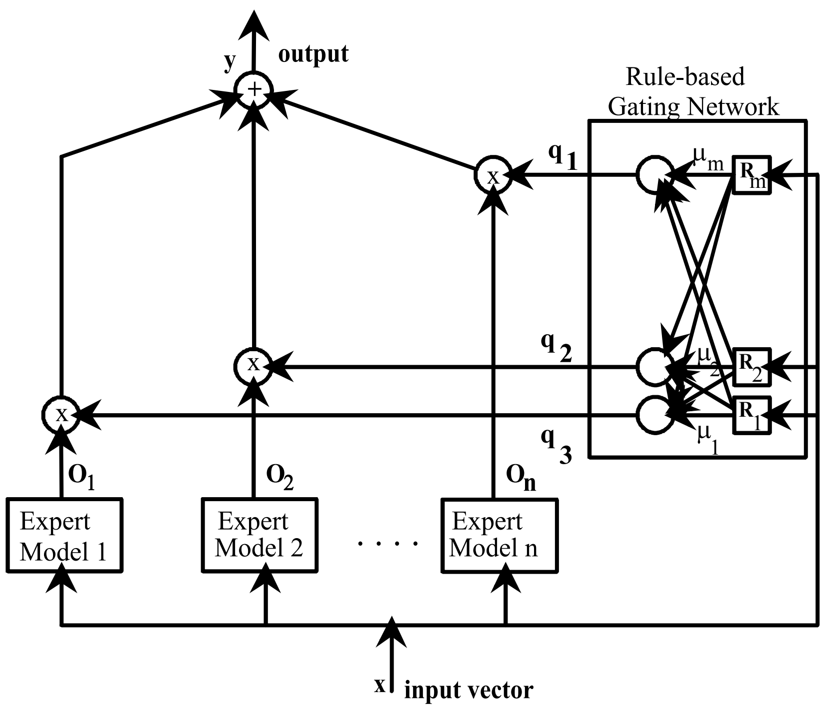

Figure 4: Fuzzy mixture of expert network

4. MIXTURE OF EXPERTS FOR BUILDING DESIGN SUPPORT

The ME structure is equivalent to a RBF network , in which each radial basis center represents a receptive with the normalized gaussian

$$q_j(x) = \frac{\mu_j}{\sum_{i=1}^{n} \mu_i} \quad \textit{and} \quad \mu_i = \exp(-\| x - x_o \|^2 / 2\sigma_i^2) \tag{7}$$

where x_o is the mean and σ_I is the variance of the i-th gaussian. The output y is computed as the weighted sum of the activation values o_i .

$$y(x) = \sum_{i=1}^{i=n} o_i q_i(x) \tag{8}$$

where o_i is the height of the i-th gaussian at the RBF network. This is a basic description of the fuzzy RBF network as mentioned earlier and it is the same description given by eqs.3 and 4 for ME. That is, a fuzzy RBF network becomes a ME system with radial basis gating function.

The basic two-level ME structure is used in this building design research using ME network, where the system consists of two RBF networks as expert networks. At the same time, the gating network is also represented by a RBF network. The radial basis functions used are gaussian functions, which play the role of membership functions in this neuro-fuzzy system. The training of the network is performed by optimization using minimum squared-error criterion. The building data used is obtained from an apartment building in Amsterdam and described earlier (Ciftcioglu et al.1998, 2000) in connection with another research on soft computing in building design using RBF network trained by Orthogonal Least Squares (OLS) method (Chen et al., 1991). The data involves an input-output structure with six inputs and four outputs for 23 cases. In the earlier study, the data were used for training of a RBF network for estimation of some test cases for building decision support. Here, from the viewpoint of performance evaluation of the ME network against the performance of a feedforward neural network, the comparison of the training results are of interest. Accordingly, the comparison results only are presented and the performance and merits of the ME network is identified in this limited scope. The training results of a RBF network with orthogonal least squares with 8 hidden nodes are presented in Fig.5 (a) and the training results of a ME network with two experts each having 4 hidden nodes is presented in figure 5(b). The comparison of the training results indicates that the ME network presents slightly better results than a single RBF network. However, it uses an additional RBF network as the gating network, although the total number of hidden nodes is the same in both cases. The better performance can be attributed to this added complexity and optimal performance of the gating network.

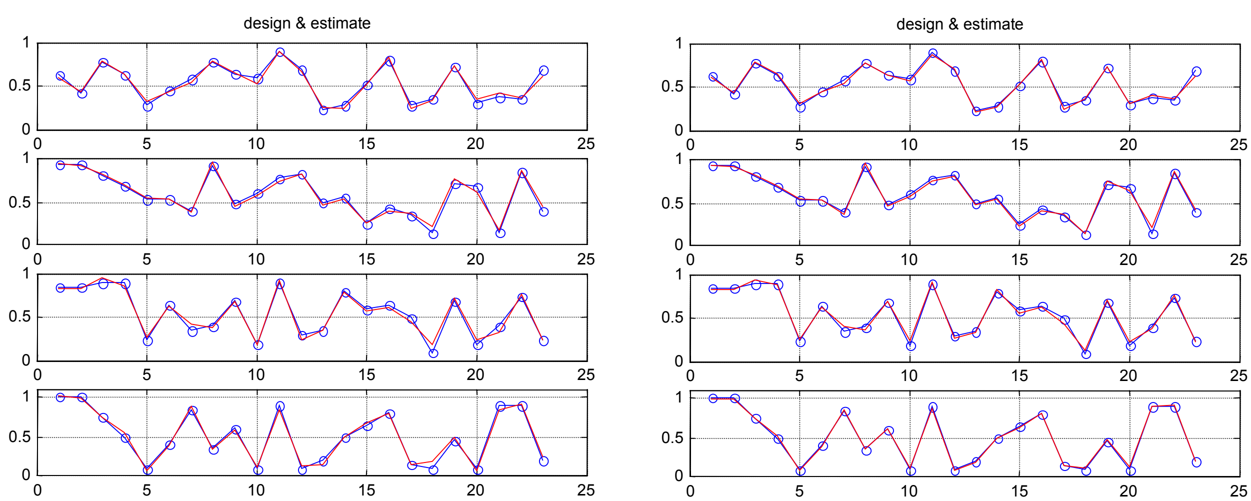

Figure 5: Training results for bulding design data by neuro-fuzzy approach. (a) A RBF network with 8 hidden nodes. (b) ME network with two experts. The true values are indicated by circles

5. DISCUSSION AND CONCLUSIONS

Although a multi layer perceptron (MLP) neural network can be trained to learn and generalize from training data, it may need an excessive number of hidden nodes so that the net result is a bulky network which in turn makes the convergence of the training process difficult. Therefore, the basic idea to tackle a complex problem by dividing it into simpler problems whose solutions can be combined to yield the final solution is rather appealing. The ME architecture using this principle weights the input space by using the posterior probabilities that expert networks generate for getting the output from the input. The training algorithm EM makes the original complicated statistical maximum likelihood (ML) problem decomposed into several ML problems. Apparently, the performance of an EM algorithm highly depends upon solutions to separate maximization problems. To tackle these problems Jordan and Jacobs (1994) propose an iteratively re-weighted least squares (IRLS) algorithm which has instabilities in multi-class classification. Therefore improved learning algorithms for ME networks are studied in the literature (Chen et al., 1999 ; Xu, 1998). Also, the normalisation ,e.g., softmax function (Bridle, 1991) made at the output of the gating networks due to probabilistic considerations have several side effects (Berthold and Diamond, 1998). Referring to these studies, it is rather straightforward to think of the possibilities of ME systems to integrate fuzzy inference systems into them. Generally, expert neural network structures, such as MLP networks, are non-parametric black boxes, that is, the parameters have no physical meaning. In contrast, the parameters in fuzzy logic systems have physical meaning. Unlike the pure NN approach, fuzzy systems require a linguistic rule-base which may be uncomplete. This implies that in the ME structure we can implement NN arhitecture as well as fuzzy inference systems where they can be cooperating for optimal outcomes. Even, unsupervised learning or clustering rules can be seen as an extension of traditional statistical clustering and parameter estimation techniques. In such fuzzy NN structures, the EM algorithm may be a convenient tool playing the role fuzzy classifier to achieve maximum likelihood estimation for each class conditional likelihood density.

The main conclusion from this research can be drawn that, the fuzzy ME networks provides new possibilities for effective use of these networks especially when the information to be processed includes linguistic qualities. This is especially the case in building technology. The outcomes are based on some logic human-like reasoning so that the expert networks, to a certain extent, play the role of human experts engaged in building technology. Since in fuzzy ME systems, the information available can be directly incorporated and used for building up the structure, by choosing proper fuzzy logic rules and membership functions, the outcomes for decision-making can be made substantially more reliable rather than pure probabilistic interpretations. However, for pure engineering applications, the linguistic information may play minimal role in information processing by ME networks, and therefore pure probabilistic implementation and interpretations can be justified in this case. The research presented here is meant primarily as a preliminary investigation of fuzzy ME networks in building technology, to demonstrate its ample potential subject to exploration.

REFERENCES

Chen S, Cowan C.F.N. and Grant P.M., Orthogonal Least Squares Algorithm for Radial Basis Function Networks, *IEEE Trans. on Neural networks, Vol.2, No.2, March 1991*

Berthold M.R. and Diamond Jay, (1998), Constructive training of probabilistic neural networks, *Neurocomputing 19, pp.167-183*

Bridle J.S. (1991), Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition, *NATO ASI Series, Vol.F.68, Neurocomputing*

Chen K., Xu L. and Chi H. (1999), Improved learning algorithms for mixture of experts in multiclass classification, *Neural Networks 12, pp.1229-1252*

Ciftcioglu Ö, Sariyildiz S. and Veer P. v.d., 1998 , Information Ordering for decision support in building design, *D & DSS, Design & Decision Support Systems 4th International Conference on Design and Decision Support Systems in Architecture and Urban Planning, Castle Vaeshartelt, Maastricht, July 26-29, The Netherlands*

Ciftcioglu Ö., Durmisevic S., Durmisevic E. and Sariyildiz S., 1999, Artificial intelligence in building design, *ISAMA 99, Sen Sabastian 7-11 June, Spain*

Ciftcioglu Ö., Durmisevic S., Durmisevic E. and Sariyildiz S., 2000, Building design support by soft computing*, Artificial Intelligence in Design, AID-2000, 26-29 June 2000, Boston, USA*

Jordan M.I. and Jacobs R.A., 1994, Hierarchical Mixtures of Experts and the EM Algorithm, *Neural Computation 6, pp.181-214*

Jordan M.I., 1995, Convergence results for the EM approach to mixtures of experts architecture, *Neural Networks, Vol.8, No.9, 9pp.1409-1431*

Xu L. (1998), RBF nets, mixture of experts, Bayesian Ying-Yang learning, *Neurocomputing 19 (1998) pp.223-257*

Zadeh L., (1965), Fuzzy sets. *Information and Control* **8**, pp.338-353

Zadeh L., An outline of a new approach to the analysis of complex systems and decision processes*, IEEE Trans. Systems Man and Cybernetics 1973; SMC-3: 29-44*