Comparison of Selection Method of a Membership Function for Fuzzy Neural Networks

Author Details: **Md. Musa Khan**[1]
[1]PhD Fellow, Anadolu University, Faculty of Science Department of Statistics, Eskisehir, Turkey

Abstract

Fuzzy neural networks are learning machine that realize the parameters of a fuzzy system (i.e., fuzzy sets, fuzzy rules) by exploiting approximation techniques from neural networks. In this paper, we tend to illustrate a general methodology, based on statistical analysis of the training data, for the choice of fuzzy membership functions to be utilized in reference to fuzzy neural networks. Fuzzy neural networks give for the extraction of fuzzy rules for from artificial neural network architectures. First, the technique is represented and so illustrated utilizing two experimental examinations for determining the alternate approach of the fuzzy neural network.

Keywords: *Fuzzy, Neural networks, Membership function, Training data.*

1.Introduction

A neural network (NN) consists of the many artificial neurons that are connected along in line with a selected network architecture. the objective of the NN is to remodel the inputs into significant outputs. NNs contains sets of adaptive weights, i.e., numerical parameters that are tuned by an algorithmic learning rule. NNs are capable of approximating non-linear functions of their inputs. fuzzy logic may be a sort of many-valued logic. It deals with reasoning that is approximate instead of fastened and exact. In binary sets, variables might take on true (1) or false (0) values. Fuzzy logic variables might have a truth value that ranges in degree between 0 and 1. A fuzzy neural network combinations components of fuzzy and neural network computations into one connectionist architecture [1,2,3,4]. Generally, two different types of neural-fuzzy hybrids have occurred: one that uses neural networks to derive the parameters of a fuzzy system, and another that gives an implementation of a fuzzy system inside a neural network architecture. NNs and fuzzy systems may be used for solving a problem if there doesn't exist any mathematical model of the given problem. NNs and fuzzy systems entirely do have sure disadvantages and benefits. Disadvantages almost completely disappear by combining each idea and new network known as a fuzzy neural network (FNN). An FNN supported an underlying fuzzy system is trained by suggests that of a data-driven learning method derived from NN theory. FNN may be described as a collection of fuzzy rules at any time of the learning process. An FNN approximates an n-dimensional unknown function that is partially described by training examples.

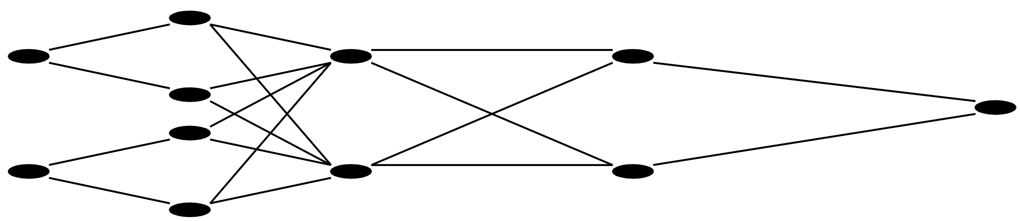
Neural-fuzzy systems may be accustomed acquire solutions involving the neural network's benefits of model-free learning, good generalization, and powerful non-linear mapping capabilities. The foremost strength derived from the fuzzy logic element may be a system which will both be initialized by the present semantic data and have structured information (knowledge) extracted from it in an explainable format. So, the "black box" problem, that is widely considered a principle weakness of neural networks may be lessened.

In Nauck [9] definition: "A hybrid neuro-fuzzy system may be a fuzzy system that uses a learning algorithm supported gradients or impressed by the neural networks theory (heuristically learning strategies) to work out its parameters (fuzzy sets and fuzzy rules) through the patterns process (input and output)."

A neuro-fuzzy system may be interpreted as a collection of fuzzy rules. this technique may be total created from input-output data or initialized with the a priori information within the same manner of fuzzy rules. The resultant system by fusing fuzzy systems and neural networks has as benefits of learning through patterns and the straightforward interpretation of its functionality. There are a series of steps to implement a basic neural-fuzzy system: (i) convert real-valued data into a fuzzified representation; (ii) train the fuzzified information with a neural network; then (iii) de-fuzzified the result to provide real values of the specified output. when the system is trained to satisfaction, fuzzy rules may be extracted from the trained neural network.

1.1. The FNN network

We use a neural-fuzzy model, referred to as FNN [2,3] that consists of five layers: an input variable layer; a condition element (fuzzification) layer; a rule layer; an action element (de-fuzzification) layer; and an output layer. In Figure 1, every one of the two inputs may be fuzzified in the conditioning layer by representing the degree of their membership in fuzzy sets. when the FNN fuzzy neural network is trained, the backpropagation training algorithmic rule is employed to regulate either the weights of the center layers (layers 2, 3, and 4) or the weights of all the layers, as well as the fuzzification and de-fuzzification layers [11].



Input Condition element Rule Action element Output

Figure-1: The FNN architecture

2. Fuzzy Membership Functions (FMFs)

Fuzzy sets are sets whose components have degrees of membership value [0,1]. A fuzzy set is completely characterized by a membership function (MF). A membership function (MF) may be a curve that defines however every purpose within the input space is mapped to a membership value (or degree of membership) between 0 and 1. The characteristics of membership function are subjective measures and not probability functions. Generally, fuzzy memberships are generated by setting the fuzzy membership as a function of the distance between the data purpose and its category center [10]. several fuzzy membership functions are planned supported this idea. although those strategies may cope with outliers or misclassification noise within the classification problem, they weren't capable of handling attribute noise accurately. The input space is usually expressed because the universe of discourse, an elaborate name for a straightforward conception. The neurons within the input layer of FNN architecture denotes the input variables as crisp values. These values are fed to the condition component layer that performs fuzzification. Constructing the correct membership functions (MFs) may be a key within the fuzzification process. many approaches for constructing and adapting membership functions are projected [1,2,3,5]. three methods for producing the membership functions used by FNN are as following:

2.1. A fixed center-based membership approach

A fixed center-based MFs approach may be instigated by exploitation three-point triangular membership functions. The triangles are finished with the minimum and most points connected to adjacent centers, or shouldered within the case of the first and last membership functions. The membership functions are opened equally in line with the minimum and maximum values of the input data. With these triangular membership functions every input value can belong to no over two fuzzy sets, and their membership degrees can continuously total to one. although this approach is easy, the division into equally spaced membership functions could also be inaccurate and inappropriate for a few data sets.

2.2. Manual adjustment of MFs centers

The center is adjusted manually for triangle-shaped MFs. during this case a stiff partitioning is usually used to type an area among that every center will move however not cross. Generally, manual changes solely change the membership functions slightly and in some cases, it should believe human experience and knowledge.

2.3. Modified backpropagation training algorithm for adjustment of MFs

This algorithm [1,2,5,6] permits modifications to be created to MFs, subject to constraints necessary for retaining semantic which means. within the FNN structure, the fuzzification layer and the de-fuzzification layer change their input connections supported easy and intuitive formulas. These changes replicate the ideas depicted by the layers and must satisfy the constraints imposed on the membership functions.

2.4. Genetic algorithms for adaptation of MFs

Genetic algorithms (GAs) are formed based on the principles of genetic science and natural process. they are widely used as powerful search techniques. Adaptation of fuzzy membership functions has been shown to guide to more effective solutions than manual alteration. Their work applied in [8] proposed an algorithmic rule that uses two factors (shift factor and shrink factor) to form little changes to the breadth and center positions of the membership functions. an identical approach [6] has been taken within the GA module of FNN, wherever once more fixed, impulsive boundaries are found out to limit the number of attainable adjustment. during this case, only the centers ought to be described within the somatic chromosome of the GA modules, fastmoving up the adaptation process and probably reducing spurious native minima within the approach utilized by [8]. The utilization of GAs, however, may be computationally costly. in the following, we tend to describe an alternate approach based on the $\chi2$ statistic to make membership functions before the FNN structure is made and trained.

3. Materials and Methods

In this part, we discuss data and approaches which will be used to compare the selection methods of a membership function for fuzzy neural networks. We conducted two experiments for data. One is Syria diabetes data set, and another one is Faldo Golf Course in the Antalya of Turkey. Two approaches will be applied for selecting the alternate approach of the membership function for fuzzy neural networks. Experiments and approaches will be illustrated elaborately in examples and experiments section.

4. Membership function (MFs) based on the $\chi2$ Approach

The goal of $\chi2$ primarily based membership approach is to settle on the optimum membership functions via a statistically-based algorithmic rule that may build neuro-fuzzy computation a lot of efficient. The $\chi2$ based membership approach performs automatic discretization of the data, which may cause an acceptable choice of the number and widths of the membership functions.

4.1. $\chi2$ algorithm

The $\chi2$ algorithmic rule [7] may be a general algorithm the first that uses the $\chi2$ statistic to discretize numeric attributes and reach feature selection. It conducts a significance test on the connection between the value of an attribute and the classes of neighboring classifications and consists of two phases. Within the first phase, data values for a given attribute are sorted and related to an interval (initially, the number of intervals equals the number of distinct values of an attribute). Then the $\chi2$ value is adjacent intervals, and adjacent intervals are unified if their $\chi2$ values fall below a definite value. This is often finished every attribute, and the merged intervals currently represent a discretization of the data set. As intervals are merged, inconsistencies will seem (identical inputs yielding totally different outputs). The above method is continued till an inconsistency rate, is exceeded within the discretized data. The second part of the $\chi2$ approach is a finer method of the first part. every attribute takes turns for merging. If throughout this process an attribute has been merged to a single interval, then it'll not be concerned in any merging.

4.2. Constructing Membership Functions via the $\chi2$ algorithm

The merged intervals for a given attribute (input) from the $\chi2$ algorithmic rule (phase1or phase2) confirm the number and the widths of the membership functions. Four-point trapezoidal membership functions are used that cause every input value to belong to a most of two membership functions, the membership degrees that can continuously sum to one. The boundaries of each membership function are determined as follows: the smaller interval is chosen from each combine of adjacent intervals. Then the half (or quarter) size areas of those smaller intervals are calculated. The fuzzy boundaries are obtained by setting those areas on both sides of every interval boundary.

Now, illustrate how the χ2 based membership approach works exploitation the Syrian diabetes database, that contains 778 patterns. each pattern is represented exploitation eight numeric attributes: a number of times pregnant, body mass index, etc. the class value is either tested positive or negative for diabetes. the data set is split at random into two sets, with 574 patterns used for finding the correct membership functions supported the χ2 algorithmic rule and the rest of the test. Here the two stages (phase1 and phase2) of the χ2 process are shown to demonstrate the behavior of the algorithmic rule. The inconsistency rate was set at five-hitter. With the phase1 stage, the data was discretized such all eight attributes have a similar minimum significance level (Significance level = 0.99, χ2=6.60), and the number of inconsistencies was kept beneath the inconsistency threshold of 28 values (28 = 574 ×5%). The phase2 stage then proceeds till no additional attribute value merging is feasible without sacrificing discriminating power. Table1 shows the intervals and attribute, "a number of times pregnant." The results for this attribute at the ends of the two stages are shown in Table2. With the χ2 threshold 6.60, for example, four intervals (discrete values) are required for the first attribute: [0, 1) -> 1, [1, 3) -> 2, [3, 7) -> 3 and [7, +∞) - > 4. The membership functions of this input attribute can be illustrated within the next section.

Table1: The initial intervals and χ^2 values for number of times pregnant

Interval	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	17
χ^2	6.44	0.02	2.53	0.16	0.34	0.95	2.35	0.64	0.01	0.60	0.02	0.06	0.27	2.04	0.1	0.2	

Table 2: The intervals and χ2 values for attribute "number of times pregnant" after Phase1 and Phase2.The χ2 thresholds **(**a) 6.60 (b) 20.73

(a)

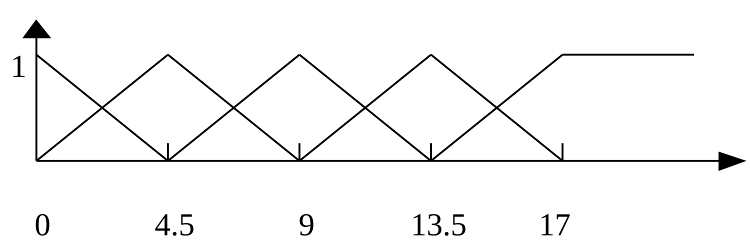
Interval	0	1	3	7
χ2	7.88	11.49	12.08	

(b)

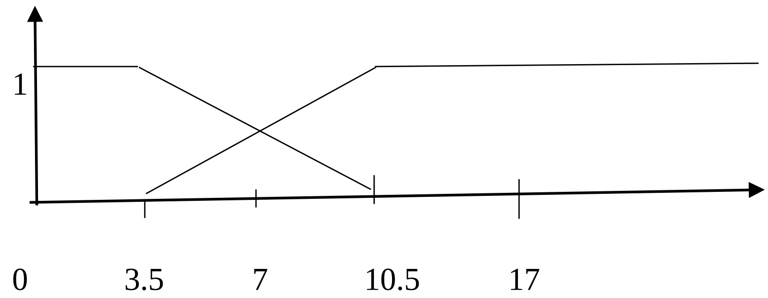
Interval	0	7
χ2	27.34	

MFs based on χ2 approach

The same number of training and testing samples as those in the previous section were taken from the data set to be used with the χ2 approach. The number of membership functions for every input variable is totally different here, as an example 2 membership functions for "number of times pregnant" and 10 membership functions for "body mass index." The membership functions "for a number of times pregnant" are shown in Figure 2(b). We tend to use the same fuzzy neural network architecture as in the previous section with 40 input nodes, 20 hidden nodes, and 5 output nodes. Its analysis is shown in Table 3(b



(a) number of time pregnant



(b) number of time pregnant

Figure 2. MFs for the number of times pregnant by (a) center-based and (b) $\chi 2$ based approach.

5. Experiments

In this section, we employ an FNN fuzzy neural network, as described in the first section, to test different training and adaptation strategies of the membership functions for two example applications. The Syrian diabetes dataset is used in the first experiment. The second experiment is based on the problem of determining suitable sites for Faldo golf courses in the Antalya of Turkey.

5.1. Syrian Diabetes Data Set

MFs based on fixed center-based approach 574 patterns were randomly chosen from the original data set for training , and the rest (204 patterns) were for testing. All eight input variables are represented as five fuzzy values. The membership functions for “number of times pregnant” are illustrated in Figure 2(a). To match the linguistic input values, a three-layer fuzzy neural network was created with 40 input nodes (each of which is associated with a fixed membership function of the input variables), and 5 output nodes for the de-fuzzification module. The backpropagation algorithm is used to train the middle layers with the 574 fuzzified samples. The test results are shown in Table 3(a).

Table 3. The test result for Syrian Diabetes data set (a) MFs produced by center-based approach (b) MFs produced by $\chi 2$ based approach

(a)

Training epochs	100
Error after training	0.1610
No difference	148 (72.55%)
One class difference	56 (27.45%)

(b)

Training epochs	100
Error after training	0.1450
No difference	158 (77.45%)
One class difference	46 (22.45%)

The performance is found to be better with the $\chi 2$ method than that with the center-based approach.

5.2 Faldo Golf Course Problem

We assume that a suitable location for a golf course in the Antalya can be determined from the observed data of mean summer temperature, mean annual rainfall, mean altitude, and distance. The output of golf course suitability is taken to have five possible values, ranging from 0 (very unsuitable) to 4 (very suitable) [4].

MFs based on fixed center-based approach

All input attributes: altitude, rainfall, temperature, and distance are represented as five fuzzy values. The membership functions for rainfall are illustrated in Figure 3(a) [Purvis et al., 1997]. A fuzzy neural network, with 20 input nodes, 20 hidden nodes, and 5 output nodes was trained, using 1000 samples of the 100000 total data examples. After evaluation over the entire data set, the FNN was found to classify 83.6% correctly, and another 16.4% were off by one membership class

MFs based on χ2 approach

MFs for golf course problem based on χ2 approach are shown in Table 4, which also lists the number of the membership function for each attribute based on the center-based approach. The membership functions for rainfall are shown in Figure 3(b) A three-layer fuzzy neural network was created with 18 input nodes, 20 hidden nodes, and 5 output nodes to calculate the output membership degrees. The fuzzy neural network was trained with the 10,0000 fuzzified samples. After the FNN was tested again over the entire data set, 93.0% correct values over the full test set and 7.0% were off by one membership class.

Table 4. *Number of MF based on two methods*

	Center-based method	χ2 method
Altitude	5	4
Rainfall	5	4
Temperature	5	5
Distance	5	4

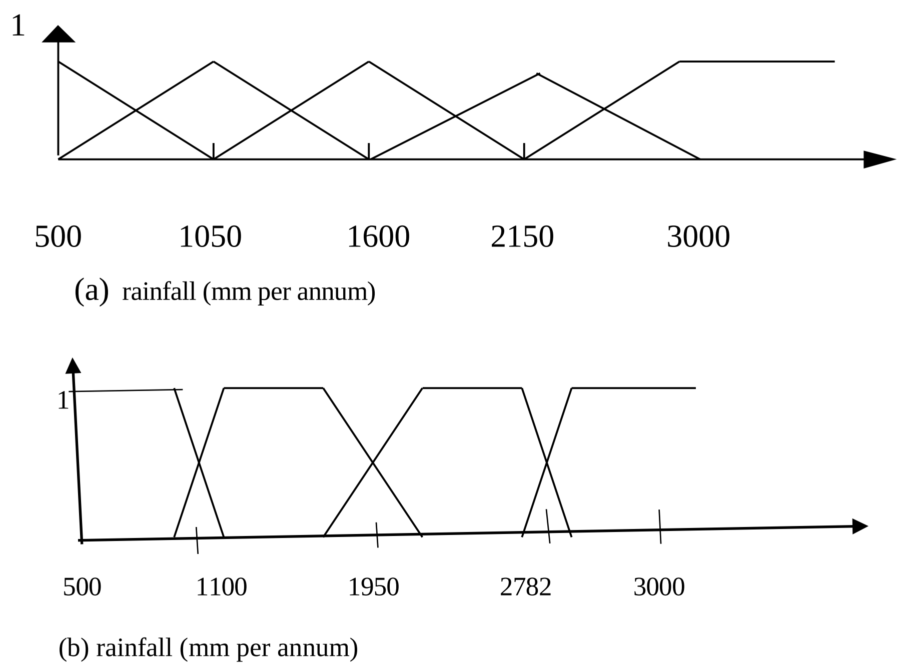


Figure 3. MFs for the rainfall by (a) center-based approach and (b) χ2 based approach

6. Conclusion

The fixed center-based approach to membership function selection is for the designer to form a subjective, typically arbitrary, determination of the number of fuzzy set values, each with a similar function width. The connectionist approach, like exploitation genetic algorithms, to select membership functions is based on training a neural network. The χ2 based membership offers an alternate approach to let the data confirm the nature of the membership functions. The results show that $\chi 2$ approach will lead to satisfactory performance for fuzzy-neural networks.

Acknowledgements: This work is funded by Science Faculty, Anadolu University.

References

i. *Horikawa et al,1992, Horikawa, A., Furuhashi, T. and Uchikawa, Y. (1992) On Fuzzy Modeling Using Fuzzy Neural Networks with the Back-Propagation Algorithm. IEEE Transactions on Neural Networks, Vol. 3, No. 5,801 – 806.*

ii. *Kasabov, 1993, Kasabov, N., (1993) Learning Fuzzy Rules and Membership Functions in Fuzzy Neural Networks, Proceeding of ANNES '93, Dunedin, New Zealand.*

iii. *Kasabov, 1996 Kasabov, N., (1996) Foundation of Neural Networks, Fuzzy Systems and Knowledge Engineering. MIT Press, Cambrige, MA.*

iv. *Purvis et al, 1997, Purvis, M., Kasabov, N., Benwell, G., Zhou, Q. and Zhang, F. (1997) Neuro-fuzzy Methods for Environmental Modeling, to appear in Proceedings of the Second International Symposium on Environmental Software Systems, Chapman and Hall, London.*

v. *Hauptmann et al, 1995, Hauptmann, W., Heesche, K. (1995) A neural Net Topology for Bidirectional Fuzzy-Neuro Transformation. Proceedings of the International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and the Second International Fuzzy Engineering Symposium, Yokohama, Japan, Vol. 3, 1071 – 1718*

vi. *Kasabov et al, 1997, Kasabov, N., Kim, J., Watts, M. and Gray, A. (1997) FuNN/2–A Fuzzy Neural Network Architecture for Adaptive Learning and Knowledge Acquisition.To appear in Information Science Application.*

vii. *Setiono, 1997 Setiono, R. (1997) Extracting Rules from Neural Networks by Pruning and Hidden-unit Splitting. Neural Computation 9, No. 1, 321 – 34.*

viii. *Mang et al, 1995, Mang, G., Lan, H. and Zhang, L. (1995) A Genetic-based method of Generating Fuzzy Rules and Membership Functions by Learning from Examples. Proceedings of ICONIP '95, China, Vol. 1, 335 – 338.*

ix. *D. Nauck, F. Klawon; R. Kruse, "Foundations of Neuro-Fuzzy Systems", J. Wiley & Sons, 1997*

x. *Lin CF, Wang SD. Fuzzy support vector machines. IEEE T Neural Networ 2002; 13: 464-471.*

xi. *Kasabov, Nikola, Qing Qing Zhou, and Martin Purvis. "A membership function selection method for fuzzy neural networks." (1997).*