Construction of Fuzzy Sets Representing Medical Concepts with an Artificial Neural Network

Gerhard Hipf, Michael Schuerz, and Klaus-Peter Adlassnig
Department of Medical Computer Sciences
Section on Medical Expert and Knowledge-Based Systems
University of Vienna Medical School
Spitalgasse 23, A-1090 Vienna, Austria
email: e9325023@stud4.tuwien.ac.at, {Michael.Schuerz|kpa}@akh-wien.ac.at

Abstract

In this paper a method for constructing fuzzy membership functions with an artificial neural network is being presented. This was carried out from data of laboratory test results from patients. Fuzzy membership functions were generated to represent vague linguistic, medical concepts for the data-to-symbol conversion unit of the medical knowledge-based system CADIAG-II.

1 Background

Efficient knowledge acquisition and representation are one of the central challenges for the successful construction and subsequent use of medical expert and knowledge-based systems in clinical practice. Since medical knowledge is immanently uncertain over wide ranges, fuzzy sets are used to deal with vague linguistic, medical concepts such as *reduced, normal, elevated, highly elevated*, and *very highly elevated*. Beginning with explorative laboratory test results from the patient data basis of the medical information system WAMIS [Grabner, 1985], knowledge acquisition of fuzzy sets is carried out in this study to yield fuzzy membership functions for the *data-to-symbol conversion* unit of the medical expert system CADIAG-II [Adlassnig *et al.*, 1985; Adlassnig, 1986].

The method of knowledge acquisition presented in this paper can be considered a semi-automatically data based approach. This means that in variation to a pure data based approach, where knowledge is acquired only from data of a specific medical field, a close cooperation of a physician and a knowledge engineer is necessary. This comprises not only that serologically verified clinical diagnoses are made by a physician but also the verification of obtained knowledge in form of fuzzy sets concerning its clinical circumstances.

During the consultation process with a medical expert system obtained measured and observed laboratory test results are then associated with fuzzy compatibility values reaching from zero to unity by consideration of the respective vague linguistic, medical concepts in the *data-to-symbol conversion* unit. These findings are used to infer diagnoses with knowledge contained in a knowledge base, which contains fuzzy relationships for the *frequency of occurrence* of findings with diseases and the *strength of confirmation* of findings for diseases [Adlassnig and Kolarz, 1986].

The aim of this paper was to generate the vague linguistic, medical concept *normal* with an artificial neural network. The fuzzy sets for the vague linguistic, medical concepts *reduced, elevated, highly elevated*, and *very highly elevated* were calculated with the method described in [Schuerz *et al.*, 1999].

2 Objectives

The goal of this study was to construct fuzzy sets with an artificial neural network in the form of S- and π-fuzzy membership functions automatically from data of one reference group of healthy individuals and six disease groups of patients suffering from inflammatory liver diseases (= hepatitides) for every observed laboratory parameter for the vague linguistic, medical concepts *reduced, normal, elevated, highly elevated*, and *very highly elevated.*

3 Material and Methods

For this retrospective study, 501 case records of adult patients who suffered from inflammatory liver diseases were selected from the online medical information system WAMIS. The patients' clinical diagnoses made are listed below; each clinical diagnosis was serologically verified and thus considered to be a gold standard. The numbers in brackets are the appropriate ICD-9 codes:

- 36 patients with type A hepatitis (070.1),
- 114 patients with type B hepatitis (070.3),
- 22 patients with non A non B hepatitis (070.5),
- 269 patients with chronic hepatitis (571.4, 571.40, 571.41, 571.42, 571.48, 571.49),
- 27 patients with alcoholic hepatitis (571.1),
- 33 patients with hepatitis carriers (V02.6).

167 case records of patients with psychophysiologic disorders (306.9) formed the reference group.

The patients above received stationary treatment in the Vienna General Hospital (AKH-Wien) within the period from April 1, 1976 to March 31, 1986. Due to the given technical and organizational setting, more recent data could not be obtained. Consequently, no distinction between hepatitis C, D, E, F, and G was made; the general diagnosis "hepatitis non A non B" was used instead. Based on the clinical picture and the prevalence's of the different types of hepatitides in Vienna, most patients who suffered from the disease "hepatitis non A non B" would nowadays be regarded as type C hepatitis cases.

The following parameters for each patient were investigated:

- alanine aminotransferase,
- alkaline phosphatase,
- aspartate aminotransferase,
- bilirubin,
- gamma-glutamyltranspeptidase,
- lactate dehydrogenase, and

the electrophoresis parameters

- albumin,
- alpha 1 globulin,
- alpha 2 globulin,
- beta globulin, and
- gamma globulin.

These parameters are important with patients presenting inflammatory liver diseases. Based upon their patterns, they can give clinical evidence for the presence of a particular type of hepatitis.

From this data material fuzzy membership functions were constructed semi-automatically for *reduced*, *normal*, *elevated*, *highly elevated*, and *very highly elevated* to be used for the data-to-symbol conversion of laboratory test results into the respective vague linguistic, medical concepts.

3.1 The Generation of Fuzzy Membership Functions

Before determining fuzzy membership functions for the medical concept *normal* with an artificial neural network it is necessary to determine which network model to use. After an examination of different artificial neural network structures for the generation of fuzzy membership functions, two types artificial neural networks were identified to be useful for this problem, the *multilayer perceptron network* (MLP network) and the *radial base function network* (RBF network).

The MLP network is a *feed forward network*. With this artificial neural network, neurons are ordered in layers and every output of a neuron in a layer is connected to every neuron in the next layer. Connections are only from the *input layer* in direction to the *output layer*. For training a MLP network the *back propagation learning algorithm* is used [Zimmermann, 1985a]. This learning algorithm carries out a calculation of the differences of the determined output values y_i of the artificial neural network with the expected explorative output values y_i (debit-output value). These differences are used for the adjustment of the weights of the artificial neural network, starting from the output layer in direction to the input layer.

In an artificial neural network the *activation function* determines the behavior of a neuron. In MLP networks the activation function of the neurons in the input layer is a linear function; for all other neurons this is the sigmoid function.

The RBF network is also a feed forward network with a back propagation learning algorithm. But in contrast to a MLP network the activation function of the neurons in the *hidden layer* is a symmetrical function (e.g., a bell-shaped function). Hidden layers are layers between the input- and the output layer. With this type of artificial neural network the neurons in the input- and output layer have a linear activation function [Zimmermann, 1985b].

Finally, the MLP network with small changes in the structure was considered to be most suitable. For the construction of fuzzy membership functions the linear activation function in the output layer of the RBF network is more useful than the normally used sigmoid function of the MLP network. The reason therefore is that the π-fuzzy membership function constructed with the sigmoid functions of two neurons in the hidden layer needs only a final linear adjustment but no additional exponential transition. Thus, this artificial neural network type can be called a *linear multilayer perceptron network* (LMLP network). Additionally, every neuron of the artificial neural network in the hidden- and in the output layer has an extra input b (*bias*) for adjusting the neurons. With this application the optimal bias setting was found to be one.

The setup of the used artificial neural network consists of *three layers*. This means that there is only one hidden layer in the artificial neural network—the artificial neural network thus consists of one input-, one hidden-, and one output layer. The number of input- and output neurons is one, because there is only one signal x_1 on the input side, i.e., the quantitative measurement or observed value, and one signal on the output side, i.e., the frequency of occurrence of the measurement value. The hidden layer comprises two neurons since π-fuzzy membership functions are composed of two sigmoid functions. A test of the artificial neural network with more hidden neurons revealed that the complexity of the artificial neural network would increase but the quality of the results would decrease. The values w_{ij} are the weights for the input signal of the hidden-, and output neurons. Figure 1 shows the model of the used artificial neural network.

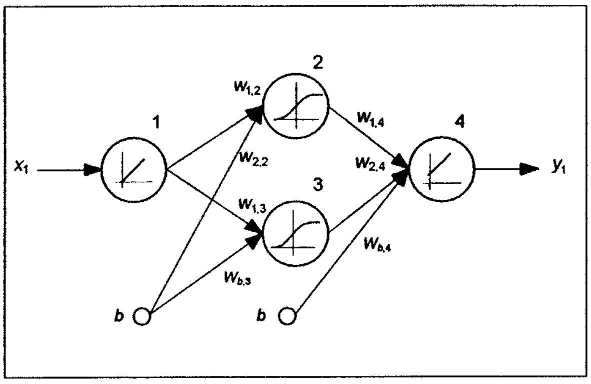

Figure 1: Structure of the used artificial neural network. The weights w_{ij} are initialized with random values prior to the training cycle whereas a fixed bias b tunes the respective neurons during training; x_1 is the input and y_1 the corresponding output value.

3.2 Properties of a Neuron

A neuron consists of an *input-*, an *activation-*, and an *output function* [Köhle, 1990]. The input function of the jth neuron is always the sum of the products of the n neuron entry signals e_i with the corresponding weights $w_{i,j}$ where i is the index of the entry signals. Hence, the neuron input function also called *net activity* net_j results in

$$net_j = \sum_{i=1,2,3,b} w_{i,j} \cdot e_i \, .$$

The activation function a_j of a neuron depends on the layer where the neuron resides within the artificial neural network. The activation function for the hidden neurons is the sigmoid function, which is usually used in MLP networks:

$$a_j = \frac{1}{1+e^{-net_j}} \, .$$

As mentioned above, the model, which was used for this study consists of an output layer, which is different from the original MLP concept. The activation function of the neuron in the input- and output layer in this case is a linear function. This means the activation function a_j of the neuron in these layers is a simple linear mapping of the net activity net_j in the neurons.

$$a_j = net_j$$

In the applied model there are two hidden neurons with sigmoid activation functions before the output layer. Hence, the input function of the output neuron is the sum of the sigmoid functions for the artificial neural network output, which results in an approximated π-fuzzy membership function.

The output function o_j with all neurons is a simple linear mapping of the activation function.

$$o_j = a_j$$

The function for the input neuron is the product of the artificial neural network entry signal x_j with the weight on this entry. This weight is always set to one. Hence the input neuron is equal to the artificial neural network entry signal and the complete function for the output of the input neuron is

$$o_{input_j} = x_j \, .$$

The complete function for the output of the hidden neurons is

$$o_{hidden_j} = \frac{1}{1+\exp^{-(\sum_{i=1,b} w_{i,hidden_j} \cdot e_i)}}$$

and for the output neuron

$$y_j = \sum_{i=2,3,b} w_{i,output_j} \cdot e_i \, .$$

These simple functions of every neuron are the basic elements of the *net function*—a function that describes the whole artificial neural network.

3.3 Net Function

Since the topology of the artificial neural network structure in our case is very simple, the used artificial neural network is resulting in a simple net function $N(x)$. This was carried out by calling the output function of the output neuron which calls the output function of the neurons in the hidden layers and that of the input neuron. With this the following formula describes the artificial neural network:

$$N(x) = w_{b,4} \cdot b + \sum_{i=2}^{3} \left(w_{i,4} \cdot \frac{1}{1+e^{-(w_{1,i} \cdot x_1 + w_{b,i} \cdot b)}} \right) = y_1 \, .$$

In this formula the bias was applied in the hidden layer and in the output layer according to the MLP network concept.

4 Results

The determination of the fuzzy membership function was carried out by the *net function* $N(x)$. Reference groups of the corresponding data record were used for training the artificial neural network (e.g., the bilirubin reference group). The data were taken from the patient database of WAMIS and transformed into the range of [−1,1] before starting to train the net. The debit-output value of the training is the frequency of occurrence of the corresponding parameter value. The following figure shows the result after the training process for bilirubin.

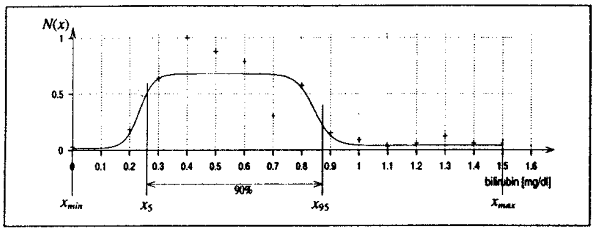

Figure 2: Resulting network function $N(x)$ after training the artificial neural network for the laboratory parameter bilirubin with data from a reference group of healthy individuals.

After training the artificial neural network, the weights w_{ij} of the artificial neural network can be put into the *net function*. In the formula below one can see the *net function* for bilirubin after training with the back propagation learning algorithm:

$$N(x) = 0{,}51 - \frac{21{,}07}{1 + e^{-(23{,}96 \cdot x - 3)}} + \frac{21{,}9}{1 + e^{-(30{,}18 \cdot x + 20{,}8)}} .$$

Fuzzy membership functions were calculated for the interval $[x_{min}, x_{max}]$, outside this interval the function value of the fuzzy membership function is equal to zero for the vague linguistic, medical concept *normal*. For defining the data points x_{min} and x_{max}, the smallest observed value of the sample of the reference group was used for x_{min} and the greatest one for x_{max}. Since the function value of the fuzzy membership function for *normal* does not have a maximum at a single position with vague linguistic, medical concepts, it is necessary to determine an interval in which the function shows the maximal membership degree unity. Since measuring errors and natural deviations influence data, the concept for *normal* can only be determined in biology in vague terms. Considering these circumstances the range for *normal* was determined to be 90 % of the area under the *net function* $N(x)$ [Richterich, 1971]. The remaining 5 % to the left and to the right of the normal range can be considered to be a gradual transition from the fuzzy compatibility value zero to unity.

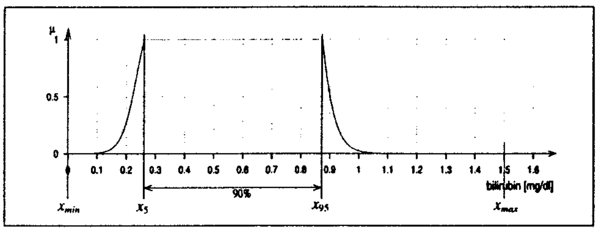

Figure 3: Fuzzy membership function for the vague linguistic, medical concept *normal* for bilirubin.

For generating the complete fuzzy membership function for the medical concept *normal*, the *net function* $N(x)$ has to be normalized to the value space [0,1]—the value space of a fuzzy set—at the intervals $[x_{min}, x_5]$ and $[x_{95}, x_{max}]$ of the abscissa. The normalized function for $[x_{min}, x_5]$ is

$$N_{low}(x) = \frac{N(x) - N(x_{Min})}{N(x_5) - N(x_{Min})};$$

analogous

$$N_{high}(x) = \frac{N(x) - N(x_{Max})}{N(x_{95}) - N(x_{Max})}$$

for the interval $[x_{95}, x_{max}]$. The complete fuzzy membership function for *normal* thus results in:

$$\mu_{\perp}(x) = \begin{cases} 0 & \text{for all} \quad x < x_{min} \\ N_{low}(x) & \text{for all} \quad x_{min} \le x < x_5 \\ 1 & \text{for all} \quad x_5 \le x \le x_{95} \\ N_{high}(x) & \text{for all} \quad x_{95} < x \le x_{max} \\ 0 & \text{for all} \quad x > x_{max} \end{cases}$$

5 Discussion

The aim of this study was to develop a method for semi-automatic generation of fuzzy membership functions using an artificial neural network.

By carrying out an abstraction step with a simple LMLP artificial neural network it was possible to approximate data of laboratory test results of sample patients. With the net function of the used artificial neural network, fuzzy membership functions could be generated for the vague linguistic, medical concept *normal* in an easy and comfortable way. The advantage of this method lies in its simplicity: A fuzzy membership function can be calculated for arbitrary data sets without changing the setting. A further advantage of this method is that artificial neural networks try to approximate a function of data during a training phase. A very good representation of the characteristics of the data can be reached with this method.

A disadvantage of this method is that learning of complex data structures with a very simple artificial neural network is not possible, but for the generation of fuzzy membership functions this artificial neural network however is sufficient. Due to the low amount of data it was not possible to test the trained artificial neural network with data other than the training data set. Thus, no quantitative value is indicating the quality of the trained artificial neural network. Instead, the evaluation of the results was made by visual comparison of the resulting function with the data.

Furthermore, a fully data based learning approach can hardly be carried out since an artificial neural network takes no knowledge about the medical field of consideration into account.

Further research will cover an evaluation of the obtained fuzzy membership functions by comparing it with different approaches like least-square-curve fitting; also an optimization of fuzzy sets with neuro-fuzzy-systems will be carried out [Bothe, 1998].

References

[Adlassnig *et al.*, 1985] K.-P. Adlassnig, G. Kolarz, and W. Scheithauer, Present State of the Medical Expert System CADIAG-2. *Methods of Information in Medicine*, 24: 13–20, 1985.

[Adlassnig, 1986] K.-P. Adlassnig, Fuzzy Set Theory in Medical Diagnosis. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-16: 260–265, 1986.

[Adlassnig and Kolarz, 1986] K.-P. Adlassnig and G. Kolarz, Representation and Semiautomatic Acquisition of Medical Knowledge in CADIAG-1 and CADIAG-2. *Computers in Biomedical Research*, 22: 361–384, 1986.

[Bothe, 1998] H.-H. Bothe, *Neuro-Fuzzy-Methoden – Einführung in Theorie und Anwendung*, Springer-Verlag, Berlin-Heidelberg-New York, pages 196–256, 1998.

[Grabner, 1985] Grabner, G., ed., *WAMIS – Wiener Allgemeines Medizinisches Informations-System*, Springer-Verlag, Berlin-Heidelberg-New York, pages 251–302, 1985.

[Köhle, 1990] M. Köhle, *Neurale Netze*, Springer-Verlag, Wien-New York, pages 62–83, 1990.

[Richterich, 1971] Richterich, R., *Klinische Chemie – Theorie and Praxis*, Karger, Basel, pages 54–55, 1971.

[Schuerz *et al.*, 1999] Schuerz, M., Adlassnig, K.-P., Lagor, C., Schneider, B., and Grabner, G., Definition of Fuzzy Sets Representing Medical Concepts and Acquisition of Fuzzy Relationships Between Them by Semi-Automatic Procedures. *Proc. of the European Symposium on Intelligent Techniques (ESIT '99)*, Crete, Greece, June 1999 (on CD-ROM).

[Zimmermann, 1985a] Zimmermann, H.-J., *Neuro+ Fuzzy – Technologien-Anwendungen*, VDI-Verlag, Düsseldorf, pages 45–71, 1985.

[Zimmermann, 1985b] Zimmermann, H.-J., *Datenanalyse – Anwendung von DataEngine mit Fuzzy Technologien und Neuronalen Netzen*, VDI-Verlag, Düsseldorf, pages 34–36, 1985.