A Novel Generic Higher-order TSK Fuzzy Model for Prediction and Applications for Medical Decision Support

Qun Song
Knowledge Engineering & Discovery Research Institute
Auckland University of Technology
Private Bag 92006, Auckland 1020 New Zealand
email qsong@aut.ac.nz

Tianmin Ma
Knowledge Engineering & Discovery Research Institute
Auckland University of Technology
Private Bag 92006, Auckland 1020 New Zealand
email mmaa@aut.ac.nz

Nikola Kasabov
Knowledge Engineering & Discovery Research Institute
Auckland University of Technology
Private Bag 92006, Auckland 1020 New Zealand
email nkasabov@aut.ac.nz

Abstract

This paper introduces a higher-order Takagi-Sugeno-Kang (TSK) neuro-fuzzy inference system and its applications in medical decision support systems. Different from most TSK fuzzy systems that utilize first-order TSK type fuzzy rules, the proposed system is composed of higher-order TKS fuzzy rules that have functions in their consequent parts of the following type: $y = b_0 x_1^{b1} x_2^{b2} \ldots x_p^{bp}$. *The type of the non-linear function has been chosen based on the rationale that it is well established in the medical practice, for example the MDRD formula and many other formulas for Glomerular Filtration Rate (GFR) evaluation used in the area of renal research and practice [8]. The proposed approach consists of three steps: (1) Apply Fuzzy C-means clustering to partition the input space; (2) Initialize a higher-order TSK fuzzy rule set according to the clustering results. For each fuzzy rule, a non-linear function of the defined type (e.g. the MDRD formula) is used in the consequent part; (3) Train the system on training data with the use of the steepest descent algorithm (Back-propagation learning algorithm) to optimize the parameters of the fuzzy rules. We illustrate the performance of the proposed system on a real case study data for GFR evaluation. For a comparison, we also present the predicted results on the same data by three other models: (i) the MDRD formula; (ii) the MLP network; and (iii) the Adaptive Neuro-Fuzzy Inference System (ANFIS) [2, 4]. The proposed system outperforms the rest as it takes into account some existing knowledge - the type of the function already developed in the past and the variables used in it, along with applying a new method for training the system.*

1. Introduction

The TSK fuzzy inference system was proposed by Takagi and Sugeno in 1985 [14]. Since then, a lot of research and applications have been developed [1, 13, 15]. The TSK fuzzy model is now established as one of the most powerful fuzzy models and has been widely and successfully used in different research areas such as: adaptive control; classification; prediction; and system identification [3, 6, 7, 10, 16].

A typical TSK fuzzy model consists of If-Then rules that have the following form:

R_l: if x_1 is F_{11} and x_2 is F_{12} and … and x_P is F_{1P},

then y is $g_1(x_1, x_2, \ldots, x_P)$ (1)

where "x_j is F_{lj}", $l = 1, 2, \ldots m$; $j = 1, 2, \ldots P$, are $m \times P$ fuzzy propositions that form m antecedents for m fuzzy rules respectively; x_j, $j = 1, 2, \ldots, P$, are antecedent variables defined over universes of discourse X_j, $j = 1, 2, \ldots, P$, and F_{lj}, $l = 1, 2, \ldots m$; $j = 1, 2, \ldots, P$ are fuzzy sets defined by their fuzzy membership functions $\mu_{Flj}: X_j \rightarrow [0, 1]$, $l = 1, 2, \ldots m$; $j = 1, 2, \ldots, P$. In the consequent parts of the fuzzy rules, y is the consequent variable, and functions g_l, $l = 1, 2, \ldots m$, are employed.

If $g_l(x_1, x_2, \ldots, x_q) = C_l$, $l = 1, 2, \ldots m$, and C_l are constants, we call this inference a zero-order TSK fuzzy inference system. The system is called a first-order TSK fuzzy inference system if $g_l(x_1, x_2, \ldots, x_q)$, $l = 1, 2, \ldots m$, are linear functions. If these functions are non-linear functions, it is called a higher-order TSK inference system [5].

For an input vector $\mathbf{x}^i = [x_1^i\ x_2^i\ \dots\ x_P^i]$ the result of the inference, y^i, or the output of the system, is the weighted average of each rule's output value indicated as follows:

$$y^i = \frac{\sum_{l=1}^{m} w_l g_l(x_1^i, x_2^i, \dots, x_P^i)}{\sum_{l=1}^{m} w_l} \quad (2)$$

where , $w_l = \prod_{j=1}^{P} F_{lj}(x_j^i);\ l = 1, 2, \dots\ m;\ \ j = 1, 2, \dots, P.$

Because the fuzzy inference systems are considered as an effective aggregation of interconnected subsystems, described with simple models, the first-order TSK models are mostly used so far. In some cases, however, certain kind of non-linear functions would be more appropriate and more effective to use in a TSK system. In our research, we use the following type of a function: $y = b_0\, x_1^{b1}\, x_2^{b2} \dots x_p^{bp}$. This type is very much used in many medical areas, such as renal function evaluation through GFR estimation with the use of the MDRD formula [8]. The function above is used as the consequent part of fuzzy rules to build a higher-order TSK fuzzy system. The system is trained on training data to optimize the parameters in both antecedent and consequent parts of the fuzzy rules.

2. The problem of accurate estimation of GFR in advanced renal impairment

The accurate evaluation of renal function is fundamental to the nephrology practice. The early detection of renal impairment will allow for the institution of appropriate diagnostic and therapeutic measures, and potentially maximise preservation of intact nephrons.

Glomerular filtration rate (GFR) [8] is traditionally considered the best overall index to determine renal function in healthy and diseased patients. Most clinicians rely upon creatinine clearance (CrCl) as a convenient and inexpensive surrogate for GFR. CrCl can be determined by either timed urine collection, or from serum creatinine using equations developed from regression analyses such as that by Cockcroft-Gault formula, but the accuracy of CrCl in terms of GFR prediction is limited by methodological imprecision and the systematic bias.

Recently, the Modification of Diet in Renal Disease (MDRD) study developed a new formula to more accurately evaluate the GFR. The formula uses the following variables: age, sex, gender, height, serum creatinine (Screat), serum albumin (Salb) and blood urea nitrogen concentrations (Surea) and is defined as follows:

$$GFR = 170 \times Screat^{-0.999} \times Age^{-0.176} \times 0.762 (\text{if sex is female}) \times 1.18 (\text{if race is black}) \times Surea^{-0.17} \times Salb^{0.318} \quad (3)$$

However, the formula can be misleading as to the presence and progression of renal disease in new patients. Therefore a method to improve the estimation of GFR in advanced renal impairment is to be investigated.

Using a higher-order TSK fuzzy model, here we obtain more accurate results for the GFR prediction than with the use of the MDRD formula. In the next section a novel learning algorithm of higher-order TSK fuzzy model is described. In section 4, the model is applied to real data for GFR prediction and the result is compared with some results obtained by other models.

3. A generic higher-order TSK Fuzzy model

3.1. Model Initialization

To partition the input space for creating fuzzy rules and obtain initial parameter values of the fuzzy rules, a *Fuzzy C-means clustering* method [2] is applied. The cluster centres are taken as initial values of the centres of *Gaussian* fuzzy membership functions and in each cluster, the maximum distance between the cluster centre and samples is taken as the initial value of the width of the corresponding *Gaussian* membership function.

As the consequent part of each higher-order TSK fuzzy rule, the non-linear function is a MDRD formula like function that can be expressed in the following form:

$$y = b_0 x_1^{b_1} x_2^{b_2} \cdots x_p^{b_p} \quad (4)$$

We apply a logarithm transformation to Eq. (4) and obtain Eq.(5):

$$\ln(y) = b_0 + b_1 \ln(x_1) + \dots + b_p \ln(x_p) \quad (5)$$

We then use *least-square estimator* (LSE) on the samples in each cluster to obtain initial values of the parameters b_j , $j = 0, 1, 2, \dots$, P, where P is the number of inputs.

3.2 Learning and parameter optimization through the steepest descent algorithm

In the higher-order TSK fuzzy model training algorithm, described below, the following indexing denotation is used:

- Training data pairs: $i = 1, 2, \dots, N$;
- Input variables: $j = 1, 2, \dots, P$;
- Fuzzy rules: $l = 1, 2, \dots, M$;
- Training epochs: $k = 1, 2, \dots$.

The training procedure consists of the following steps. Consider the system is having *P* inputs and one output. Suppose that it has *M* fuzzy rules, where the *l*-th rule has the form of

R_l: if x_1 is F_{11} and x_2 is F_{12} and … and x_P is F_{1P},

then $y_l = b_{l0} x_1^{b_{l1}} x_2^{b_{l2}} \cdots x_p^{b_{lp}}$

$$l = 1, 2, \ldots, M \tag{6}$$

Here, F_{lj} are fuzzy sets defined by the following *Gaussian* type membership function

$$GaussianMF = \alpha \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \tag{7}$$

When the model is given an input-output pair ($\boldsymbol{x}_i$, d_i), it calculates the following output value:

$$f(\boldsymbol{x}_i) = \frac{\sum_{l=1}^{M} y_l \prod_{j=1}^{P} \exp\left[-\frac{(x_{ij}-m_{lj})^2}{2\sigma_{lj}^2}\right]}{\sum_{l=1}^{M} \prod_{j=1}^{P} \exp\left[-\frac{(x_{ij}-m_{lj})^2}{2\sigma_{lj}^2}\right]} \tag{8}$$

The goal is to design the system from (8) so that the following objective function is minimized:

$$E = \frac{1}{2}\left(f(\boldsymbol{x}_i) - d_i\right)^2 \tag{9}$$

Below, the steepest descent algorithm (BP) [5, 9, 15,] is applied to optimize the parameters b_{lj}, m_{lj}, α_{lj} and σ_{lj} in the fuzzy model:

$$b_{l0}(k+1) = b_{l0}(k) - \frac{\eta_b}{b_{l0}(k)} y_l(k)[f(\boldsymbol{x}_i) - d_i]\Phi_l(\boldsymbol{x}_i) \tag{10}$$

$$b_{lj}(k+1) = b_{lj}(k) - \eta_b y_l(k) \ln(x_{ij})[f(\boldsymbol{x}_i) - d_i]\Phi_l(\boldsymbol{x}_i) \tag{11}$$

$$\alpha_{lj}(k+1) = \alpha_{lj}(k) - \frac{\eta_\alpha}{\alpha_{lj}(k)}[f(\boldsymbol{x}_i) - d_i][y_l(k) - f(\boldsymbol{x}_i)]\Phi_l(\boldsymbol{x}_i) \tag{12}$$

$$m_{lj}(k+1) = m_{lj}(k) - \eta_m [f(\boldsymbol{x}_i) - d_i][y_l(k) - f(\boldsymbol{x}_i)]\frac{[x_{ij} - m_{lj}(k)]}{\sigma_{lj}^2(k)}\Phi_l(\boldsymbol{x}_i) \tag{13}$$

$$\sigma_{lj}(k+1) = \sigma_{lj}(k) - \eta_\sigma [f(\boldsymbol{x}_i) - d_i][y_l(k) - f(\boldsymbol{x}_i)]\frac{[x_{ij} - m_{lj}(k)]^2}{\sigma_{lj}^3(k)}\Phi_l(\boldsymbol{x}_i) \tag{14}$$

where,

$$\Phi_l(\boldsymbol{x}_i) = \frac{\prod_{j=1}^{P} \alpha_{lj} \exp\left\{-\frac{[x_{ij}(k) - m_{lj}(k)]^2}{2\sigma_{lj}^2(k)}\right\}}{\sum_{l=1}^{M} \prod_{j=1}^{P} \alpha_{lj} \exp\left\{-\frac{[x_{ij}(k) - m_{lj}(k)]^2}{2\sigma_{lj}^2(k)}\right\}} \tag{15}$$

η_b, η_m, η_α and η_σ are learning rates for updating the parameters b_{lj}, m_{lj}, α_{lj} and σ_{lj} respectively.

4. Experimental results for the GFR prediction and a comparative analysis with other techniques

A real data set is used here for experimental analysis. The data set has 447 samples, collected at hospitals in New Zealand and Australia. Each of the records includes 6 variables (inputs): age, sex, gender, serum creatinine, serum albumin, and blood urea nitrogen concentrations, and one output - the Glomerular Filtration Rate (GFR) value. All experimental results reported here are based on 10 experiments with the same model and parameters and the results are averaged. In each experiment 70% of the whole data set is randomly selected as training data and another 30% as testing data. The results are listed in Table1. They include the number of fuzzy rules (for ANFIS and for the proposed TSK fuzzy model) or neurons in the hidden layer (for the MLP model), the test RMSE (root mean square error), and the test AAE (average absolute error). Figure.1 shows spot-graphs for each prediction result by different models.

Table 1. Experimental Results with Different Models

Model	Rules or Neurons	RMSE	AAE
MDRD	-	5.88	7.74
MLP	16	5.75	8.44
ANFIS	27	5.48	7.49
Higher-order TSK Model	12	5.08	6.97

5. Conclusions and directions for further research

The paper presents a higher-order TSK fuzzy model for prediction tasks. The higher- order function is of the type of: $y = b_0\, x_1^{b1}\, x_2^{b2} \ldots x_p^{bp}$ which is used in the medical practice, and especially for the prediction of the Glomerular Filtration Rate (GFR) in many renal clinics all over the world. The proposed model demonstrates superiority when compared with the widely popular MDRD formula [8], the MLP neural network model, and the ANFIS neuro-fuzzy model.

Further directions for the research include: (1) self-optimization of the learning parameters of the introduced TSK model, such as learning rate, number of fuzzy rules, etc.; (2) optimal selection of input variables for better prediction results

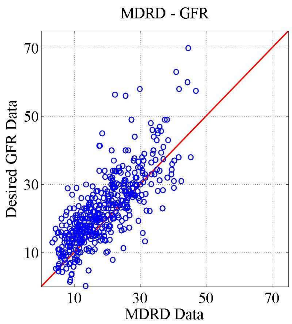

Figure1a. MDRD vs GFR

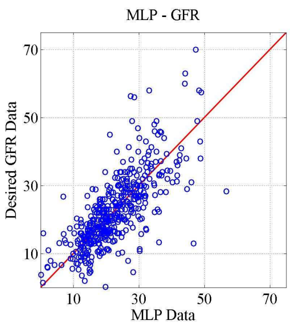

Figure1b. MLP vs GFR

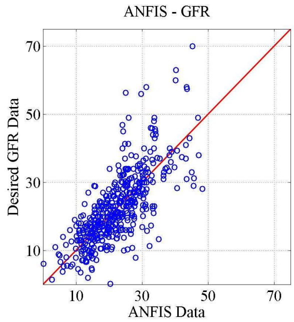

Figure1c. ANFIS vs GFR

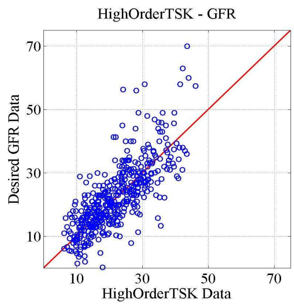

Figure1d. Higher-order TSK model vs GFR

6. Acknowledgements

This research is part of a research programme funded by the New Zealand Foundation for Research Science and Technology, contract NERF-AUT-02001. The GFR data used in the paper includes New Zealand and Australia data, collected by AUS-14 Study Group and Middlemore hospital with the participation of Dr Mark Marshall.

References

[1] Chuang, C.C., Su, S.F. and Chen, S.S., "Robust TSK Fuzzy Modelling for Function Approximation With Out-Liers", *IEEE Trans. on Fuzzy Systems*, vol. 9, pp. 810 – 821, 2001.

[2] *Fuzzy Logic Toolbox User's Guide*, The Math Works Inc., ver. 2.01, 2000.

[3] Guerra, M.T. and Vermeiren, L., "Control Laws for Takagi-Sugeno Fuzzy models", *Fuzzy Sets and Systems*, vol.120, pp. 95-108, 2001.

[4] Jang, J.S.R., "ANFIS: adaptive network-based fuzzy inference system", *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 23, pp. 665-685, 1993.

[5] Jang, J.S.R., Sun, C. and Mizutani, E., *Neuro-Fuzzy and Soft Computing*, Prentice Hall, Inc., 1997.

[6] Juang, C.F., “A TSK-Type Recurrent Fuzzy Network for Dynamic Systems Processing by Neural Network and Genetic Algorithms”, *IEEE Trans. on Fuzzy Systems*, vol. 10, pp. 155 – 170, 2002.

[7] Kasabov, N. and Song, Q., "DENFIS: Dynamic, evolving neural-fuzzy inference systems and its application for time-series prediction," *IEEE Trans. on Fuzzy Systems*, vol. 10, pp. 144-154, 2002.

[8] Levey, A.S., Bosch, J.P., Lewis, J.B., Greene, T., Rogers, N., Roth, D., for the Modification of Diet in Renal Disease Study Group, “ A More Accurate Method To Estimate Glomerular Filtration Rate from Serum Creatinine: A New Prediction Equation”, *Annals of Internal Medicine*, vol. 130, pp.461-470, 1999.

[9] Lin, C.T. and Lee, C.S.G., *Neuro Fuzzy Systems*,. Prentice Hall (1996).

[10] Mastorocostas, P.A., Theocharis, J.B. and Petridis, V.S., “A Constrained Orthogonal Least-squares Method for Generating TSK Fuzzy Models: Application to Short-term Load Forecasting”, *Fuzzy Sets and Systems*, vol. 118, pp. 215 – 233, 2001.

[11] Mendel, J.M., *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*. Prentice Hall, 2001.

[12] *Neural Network Toolbox User’s Guide*, The Math Works Inc., ver. 4, 2001.

[13] Papadakis, S.E. and Theocharis, J.B., “A GA-Based Fuzzy Modelling Approach for Generating TSK Models”, *Fuzzy Sets and Systems*, vol. 131, pp. 121 – 152, 2002.

[14] Takagi, T. and Sugeno, M., “Fuzzy identification of systems and its applications to modeling and control”, *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 15, pp. 116-132, 1985.

[15] Wang, L.X., *Adaptive Fuzzy System And Control: Design and Stability Analysis*. Englewood Cliffs, NJ: Prentice Hall, 1994.

[16] Wang, L.X. and Langari, R., “Sugeno Model, Fuzzy Discretization, and the EM Algorithm”, *Fuzzy Sets and Systems*, vol. 82, pp. 279 – 288, 1996.