GENETIC TUNING OF HIERARCHICAL MODELS BASED ON EXPERIMENTAL DATA

Kifayat Mammadova, Ibrahim Abasov, Tural Ahmedov, Oqtay Safaraliyev

Azerbaijan State Oil and Industry University
ka.mamedova@yandex.ru
abasov_i99@mail.ru
oqtay.23@mail.ru

Abstract

The multidirectional search method is applied to the individual elements of the experimental data, forming and observing potential solution sets. The aim of this study is to genetically tune linguistic models using experimental data based on a fuzzy knowledge base. Experiential knowledge is derived from experience or observation. This process manifests itself in many different situations. It can be observed and predicted using qualitative characteristics. However, the process is not always understandable in terms of fundamental principles. For the purpose of forming primary information, the article establishes a fuzzy model that determines the linguistic values of variables and the formation of their membership functions. A knowledge matrix is drawn up according to the rules connecting the input and output of the identified object, and knowledge is determined based on logical considerations of the form 'IF-THEN-OTHERWISE'. Based on the matrix, the values of the input variables x_i are associated with one of the possible types of solution d_j. This problem is solved using fuzzy logic equations. These equations are based on a knowledge base or a system of logical judgements that is similar to it. It also allowed the membership functions of various solutions to be calculated at fixed input data values for the object. Typically, a solution with a high membership function value is considered the desired solution.

Keywords: formulation of membership functions, fuzzy knowledge base, fuzzy model, linguistic variables, fuzzy knowledge base, linguistic interpretation, genetic algorithm

I. Introduction

The identification of nonlinear dynamic objects based on fuzzy logic relies on generating a fuzzy knowledge base from experimental data that links input and output variables. The method of modifying and processing experimental information in a fuzzy knowledge base is widely applied as an efficient approach for data processing in fields such as medicine, banking, management, and others. In this case, qualitative criteria will be close to the results of linguistic approximation and corresponding experimental data based on certain regularities.

In conditions of uncertainty, the information obtained about an object is incomplete, imprecise, and fuzzy. This makes tuning and managing models based on experimental data more challenging. Therefore, there arises a necessity to create and tune linguistic models based on experimental data, addressing this as an actual problem.

The goal is to find an optimal solution by genetically tuning linguistic models with experimental data, based on a fuzzy knowledge base.

Learning from experimental datasets is a hot topic in machine learning research [3,4]. This problem is characterized by a class distribution where the number of samples in one class is greater than the number of samples in the other class. Medical diagnostics, fraud detection, finance, risk management, network intrusion, etc. including, but not limited to, many real-world problems are dominated by the presence of unbalanced data sets. Furthermore, the positive or minority class is usually the class of highest interest from a learning point of view, and it is also costly if not well classified [1,5].

There are a number of intrinsic properties of data that reduce the effectiveness of learning. Some of these deficiencies within the data include the presence of small errors [6] and the presence of overlapping samples between classes.

Experimental knowledge obtained from experience or observation can be easily expressed and described in terms of qualitative conditions. The process can be observed and predicted using qualitative characteristics. However, it is not always comprehensible in terms of fundamental principles. The proposed algorithm uses the idea of identifying linguistic terms based on the maximum of the membership function, which is generalized across the entire knowledge matrix. The study of the fuzzy system based on experimental data has shown it to be more effective compared to traditional methods.

II. Problem Statement

Rules for constructing hierarchical models. When there are many input variables x_i constructing a fuzzy knowledge base about the unknown dependency (1) becomes difficult. This is related to the fact that a person can simultaneously retain an understanding of about 7±2 features in their working memory. Therefore, it is advisable to classify the input variables and construct a decision tree based on them.

The hierarchical interrelationship between the input and output variables (integral indicators) corresponding to the system of relations (1-4) is depicted in the tree structure shown in Figure 1.

$$R = f_R(X, Y, \dots, Z) \quad (1)$$
$$X = f_X(x_1, x_2, \dots, x_n) \quad (2)$$
$$Y = f_Y(y_1, y_2, \dots, y_n) \quad (3)$$
$$Z = f_Z(z_1, z_2, \dots, z_n)\ , \quad (4)$$

where, $R -$ represents the output variables (integral indicators); $X, Y, Z -$ represent classes of input variables; $x_i, y_j, z_k (i = \overline{1, l}; j = \overline{1, m}; k = \overline{1, n};) -$ are input variables that belong to the classes X, Y, Z respectively.

Fig. 1 illustrates a tree that reflects the input variables in a two-level hierarchy. However, detailing can be done across various levels: class-subclass, group-subgroup, etc. Let us assume that all variables at the top of the tree are linguistic variables expressed by the following terms:

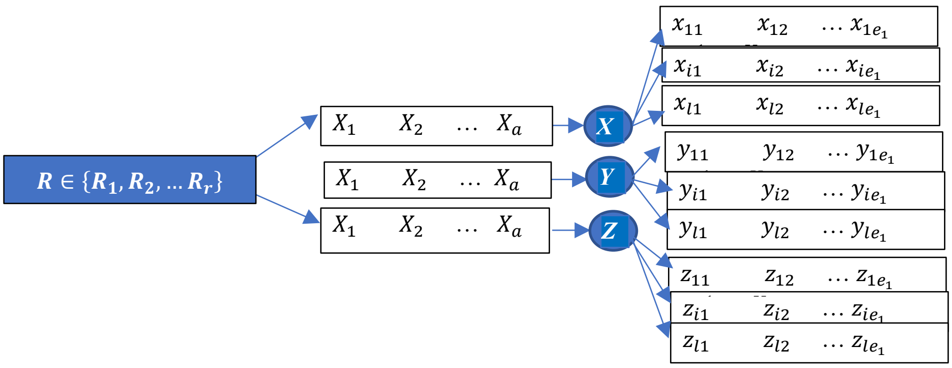

Figure 1: *Generalized Logical Inference Tree*

- $\{R_1, R_2, \dots, R_r\} -$ Set of terms for evaluating the variables R;
- $\{X_1, X_2, \dots, X_a\} -$ A set of terms for evaluating the X variables;
- $\{Y_1, Y_2, \dots, Y_b\} -$ A set of terms for evaluating Y variables;

- $\{Z_1, Z_2, \ldots, Z_c\} -$ A set of terms for evaluating the Z variables;
- $\{x_{i1}, x_{i2}, \ldots, x_{ia_i}\} -$ set of terms for evaluation of variables $x_i, i = \overline{1, l}$;
- $\{y_{j1}, y_{j2}, \ldots, y_{jb_j}\} -$ set of terms for evaluation of variables $y_j, j = \overline{1, m}$;
- $\{z_{k1}, z_{k2}, \ldots, z_{kc_k}\} -$ set of terms for evaluation of variables $z_k, k = \overline{1, n}$

The introduced sets of terms indicate the branches of the tree. Within each set, fuzzy terms are arranged sequentially from bottom to top, following the principle where the strength (i.e., the number of terms) of all sets used for evaluating the linguistic variables involved in the relations (1)-(4) may vary.

Let us describe the terms as fuzzy sets using the concept of universal sets and the membership function.

$$R_i = \int_W \mu_{R_i}(w)/w, i = \overline{1, r}, w \in W; \quad (5)$$

$$X_i = \int_{U_X} \mu_{X_i}(v_X)/v_X, i = \overline{1, a}, v_X \in U_X; \quad (6)$$

$$Y_i = \int_{U_Y} \mu_{Y_i}(v_Y)/v_Y, i = \overline{1, b}, v_Y \in U_Y; \quad (7)$$

$$Z_i = \int_{U_Z} \mu_{Z_i}(v_Z)/v_Z, i = \overline{1, c}, v_Z \in U_Z; \quad (8)$$

$$x_{ij} = \int_{U_{x_i}} \mu_{x_{ij}}(x_i)/x_i, i = \overline{1, l}, j = \overline{1, a_i}, x_i \in U_{x_i}; \quad (9)$$

$$y_{jk} = \int_{U_{y_j}} \mu_{y_{jk}}(y_j)/y_j, j = \overline{1, m}, k = \overline{1, b_j}, y_j \in U_{y_j}; \quad (10)$$

$$z_{kl} = \int_{U_{z_k}} \mu_{z_{kl}}(z_k)/z_k, k = \overline{1, n}, l = \overline{1, c_k}, z_k \in U_{z_k} \quad (11)$$

Here, W- is the universal set of the variables R, more precisely, $R_i \subset W, i = \overline{1, r} -$ is a universal set whose variables are given

$U_X, U_Y, U_Z - X, Y, Z$ are universal sets given variables, i.e., $X_i \subset U_X$,

$Y_i \subset U_Y, Z_i \subset U_Z$; $U_{x_i}, U_{y_j}, U_{z_k} - x_i, y_j, z_k, i = \overline{1, l}; j = \overline{1, m}; k = \overline{1, n}$are universal sets given by variables;

$\mu^{\xi}(\chi) -$ is the membership function of variable X of fuzzy term ξ.

We will describe the information about the relations (1)–(4) in the form of a fuzzy knowledge base that stores logical judgments about the interactions of input and output variables. A fuzzy knowledge base about relation (1) will have the following rule:

IF $(X = x_{j1})$ AND $(Y = y_{j1})$VƏ...VƏ $(Z = z_{j1})$ OR
$(X = x_{j2})$ AND $(Y = y_{j2})$ AND... AND $(Z = z_{j2})$ OR
...
$(X = x_{jh_j})$ AND $(Y = y_{jh_j})$ AND... AND $(Z = z_{jh_j})$

TNEN $R = R_j, j = \overline{1, r}$ (12)

It is possible to describe this rule in the form of M_R knowledge matrix:

$$M_R = \left| \begin{array}{cccc} X & Y & \ldots Z & R \\ x_{11} \ldots x_{1h_1} & y_{11} \ldots y_{1h_1} & z_{11} \ldots z_{1h_1} & R_1 \\ \ldots & \ldots & \ldots & \ldots \\ x_{j1} \ldots x_{jh_j} & y_{j1} \ldots y_{jh_j} & z_{j1} \ldots z_{jh_j} & \ldots R_j \\ \ldots & \ldots & \ldots & \ldots \\ x_{r1} \ldots x_{rh_r} & y_{r1} \ldots y_{rh_r} & z_{r1} \ldots z_{rh_r} & R_r \end{array} \right| \quad (13)$$

Applying the operations of union and intersection of sets (13), let us write the fuzzy knowledge base in the following form:

$$\cup_{j=1}^{h_j}\left[\left(X = x_{jp}\right) \cap \left(Y = y_{jp}\right) \cap \ldots\right] \rightarrow R_j, j = \overline{1, r} \quad (14)$$

The fuzzy knowledge base about relations (2) is defined according to the following rule:
ıf $(x_1 = {x_1}^{j1})$ and $(x_2 = {x_2}^{j1})$ and... and $(x_l = {x_l}^{j1})$ OR
$(x_1 = {x_1}^{j2})$ and $(x_2 = {x_2}^{j2})$ and... and $(x_l = {x_l}^{j2})$ OR
...
$\left(x_1 = {x_1}^{je_j}\right)$ and $\left(x_2 = {x_2}^{je_j}\right)$ and... and $\left(x_l = {x_l}^{je_j}\right)$
then $X = X_J, j = \overline{1, a}$ (15)

M_X matrices are constructed analogously to the knowledge matrix M_R described in (13). Applying the operations of union and intersection of sets (15), let us write the fuzzy knowledge base in the following form:

$$\cup_{p=1}^{e_j}[\cap_{i=1}^{m}(x_i = {x_i}^{jp})] \rightarrow X = X_j, j = \overline{1, a} \quad (16)$$

The fuzzy knowledge base about relations (3) has the following form:

IF $(y_1 = {y_1}^{j1})$ and $(y_2 = {y_2}^{j1})$ and... and $(y_l = {y_l}^{j1})$ OR
$(y_1 = {y_1}^{j2})$ and $(y_2 = {y_2}^{j2})$ and... and $(y_l = {y_l}^{j2})$ OR
...
$\left(y_1 = {y_1}^{jg_j}\right)$ and $\left(y_2 = {y_2}^{jg_j}\right)$ and... and $\left(y_i = {y_i}^{jg_j}\right)$
then $Y = Y_J, j = \overline{1, b}$ (17)

M_Y matrices are constructed analogously to the knowledge matrix M_R described in (13). Applying the operations of union and intersection of sets (18), let us write the fuzzy knowledge base in the following form:

$$\cup_{p=1}^{g_j}\left[\cap_{i=1}^{m}\left(y_j = {y_j}^{jp}\right)\right] \rightarrow Y = XY_j, j = \overline{1, b} \quad (18)$$

The fuzzy knowledge base about relations (4) has the following form:
ıf $(z_1 = {z_1}^{j1})$ and $(z_2 = {z_2}^{j1})$ and... and $(z_n = {z_n}^{j1})$ OR
$(z_1 = {z_1}^{j2})$ and $(z_2 = {z_2}^{j2})$ and... and $(z_l = {z_l}^{j2})$ OR
...
$\left(z_1 = {z_1}^{jt_j}\right)$ and $\left(z_2 = {z_2}^{jt_j}\right)$ and... and $\left(z_n = {z_n}^{jt_j}\right)$
then $Z = Z_J, j = \overline{1, c}$ (19)

M_Z matrices are constructed analogously to the knowledge matrix M_R described in (13). Applying the operations of union and intersection of sets (21), let us write the fuzzy knowledge base in the following form:

$$\cup_{p=1}^{t_j}[\cap_{i=1}^{m}(z_i = {z_i}^{jp})] \rightarrow Z = Z_j, j = \overline{1, c} \quad (20)$$

Thus, a system of fuzzy logical judgments (14), (16), (18), (20) was determined, which describes the expert information about the relations (1–4) corresponding to the generalized tree of interconnected "input - output variables" .

The principle of encoding linguistic variables and their fuzzy terms proposed above for a two-level inference tree remains unchanged for a tree of any size. The structure of the tree, the selection of the terms of the linguistic variables into its branches, as well as the selection of the fuzzy knowledge base at each non-terminal vertex of the tree, are important for the construction of models of non-linear objects.

III. Solving the problem

Extracting knowledge from experimental data. The identification of nonlinear dynamic objects based on fuzzy logic is explained by the presence of IF-THEN rules that depend on the linguistic values of input-output variables. Previously, it seemed to us that IF-THEN rules were generated by experts. What can be done if there is no such expert? In this case, there is an interest in the generation of IF-THEN rules, fuzzy knowledge base generated from experimental data.

The method of changing experimental information in a fuzzy knowledge base can be an efficient method for processing data in medicine, banking, management and other fields. In this case, due to certain regularities, the quality criteria will be close to the results of linguistic approximation and corresponding experimental data.

An object with n inputs and one output (24) is given by the following properties [3,4,9]:

$$y = f(x_1, x_2, \dots, x_n) \tag{21}$$

1. Input and output change interval

$$x_i \in [\underline{x}_i, \overline{x}_i], i = \overline{1,n}; y \in [\underline{y}, \overline{y}],$$

2. $d_j (j = \overline{1,m};)$ is a class of solutions in objects with discrete outputs:

$$[\underline{y}, \overline{y}] = \underbrace{[\underline{y}, y_1]}_{d_1} \cup \underbrace{[y_1, y_2]}_{d_2} \cup \dots \cup \underbrace{[y_{j-1}, y_j]}_{d_j} \cup \dots \cup \underbrace{[y_{m-1}, \overline{y}]}_{d_m}$$

A training sampling interval in the form of M pairs of experimental "input-output" data:

$\{X_p, y_p\}$– for objects with continuous output,

$\{X_p, d_p\}$– for objects with discrete output,

where $X_p = \{x_1^p, x_2^p, \dots, x_n^p\}$– p is the input vector in the -th pair, $p = \overline{1, M}$

It is required to summarize the information about the object (1) in the form of fuzzy logical judgments:

IF $x_1 = a_1^{j1}$ and $x_2 = a_2^{j1}$ and ... and $x_n = a_n^{j1}$ (w_{j1} by weight)

or $(x_1 = a_1^{j2})$ and $(x_2 = a_2^{j2})$ and ... and $(x_n = a_n^{j2})$ (w_{j2} by weight)

... or $[(x_1 = a_1^{jk_j})$ and $(x_2 = a_2^{1kj})$ and … and $(x_n = a_n^{1kj})]$ ($w_{j_{kj}}$ by weight)

Then $$y \in d_j = [y_{j-1}, y_j], j = \overline{1,m}, \tag{22}$$

$a_i^{jp} - p = \overline{1, k_j}$In the numbered row, x_i are the linguistic terms that evaluate the input variables; $k_j - d_j, j = \overline{1,m}$ is the number of conjunction rows corresponding to the classes; w_{jp}–[0,1] is a number characterizing the weight of the jp -numbered judgments within the range.

The knowledge base (22) corresponds to the object's model (24) in the form of the following computational relations [3-6]:

$$y = \frac{\underline{y}\mu^{d_1}(y) + y_1\mu^{d_2}(y) + \dots + y_{m-1}\mu^{d_m}(y)}{\mu^{d_1}(y) + \mu^{d_2}(y) + \dots + \mu^{d_m}(y)}, \tag{23}$$

$$\mu^{d_j}(y) = \underline{max}_{p=1,k_j}\left\{w_{jp}\underline{min}_{i=1,n}[\mu^{jp}(x_i)]\right\}, \tag{24}$$

$$\mu^{jp}(x_i) = \frac{1}{1 + \left(\frac{x_i - b_i^{jp}}{c_i^{jp}}\right)}, \quad i = \overline{1,n}, j = \overline{1,m}, p = \overline{1,k_j} \tag{25}$$

Here, $\mu^{d_j}(y) -$ is the membership function of output y for class d_j;

$\mu^{jp}(x_i)$– membership function of input x_i for a_i^{jp} term;

b_i^{jp} and c_i^{jp}– x_i are parameters that debug the membership functions of the input variables. Relations (23),(24),(25) determine the model of the object written as follows [10]:

$y = F(X, W, B, C)$ – for an object with continuous output;

$\mu^{d_j}(y) = \mu^{d_j}(X, W, B, C)$ – for object with discrete output;

$X = (x_1, x_2, \dots, x_n)$ – input vector;

$W = (w_1, w_2, \dots, w_n)$ – (25) vector of weighting rules in the fuzzy knowledge base;

$B = (b_1, b_2, \dots, b_q)$ and $C = (c_1, c_2, \dots, c_q)$ – tuning parameters vector of membership functions of fuzzy terms in (26);

$N -$ total number of lines of rules, q - total number of terms; $F -$ is a vector of input-output relations corresponding to the use of relations (24)–(25).

(3.2) we give the restrictions on the size of the knowledge base by one of the following cases:

a) $N = k_1 + k_2 + \dots + k_m \leq \overline{N}$;

b) $k_1 \leq \overline{k}_1, k_2 \leq \overline{k}_2 \dots, k_m \leq \overline{k}_m$

where, $\overline{N}$ is the maximum number of conjugation lines possible in (25); $\overline{k_j} - d_j, j = \overline{1,m}$the maximum number of possible conjunction lines in the j[th] order of the class of decisions, the number of $a_i^{jp}, i = \overline{1,n}, p = \overline{1,k_j}, j = \overline{1,m}$ linguistic terms used in (25); if the composition is not known, then it is possible to interpret them based on the values of the (b_i^{jp}, c_i^{jp}) parameters of the membership function (27). Therefore, the knowledge base synthesis (25) corresponds to obtaining the parameters matrix described in table 1 [5-7].

In mathematical programming, this problem can be formulated as follows. It is necessary to find a matrix (table 1) that satisfies the restrictions on the range of variation of parameters {W,B,C} and the number of rows:

For a facility with continuous output [9-11]

$$\sum_{p=1}^{M} \left[F\left(X_p, W, B, C\right) - y_p\right]^2 = \min_{W,B,C} \tag{26}$$

$$\sum_{p=1}^{M} \left\{\sum_{j=1}^{m} \left[\mu^{d_j}\left(X_p, W, B, C\right) - \mu_p^{d_j}(y)\right]^2\right\} = \min_{W,B,C} \tag{27}$$

Where, $\mu_p^{d_j} = \begin{cases} 1, d_j = d_p \\ 0, d_j \neq d_p \end{cases}$

It is considered appropriate to use a genetic algorithm to solve this problem.

Experimental data (x∈[0,10], y∈[-0.47,0.79]) about the object $y = f(x) = e^{\frac{-x}{4}} \cdot \sin(\frac{\pi}{2} x)$ in picture 2 generates the model described.

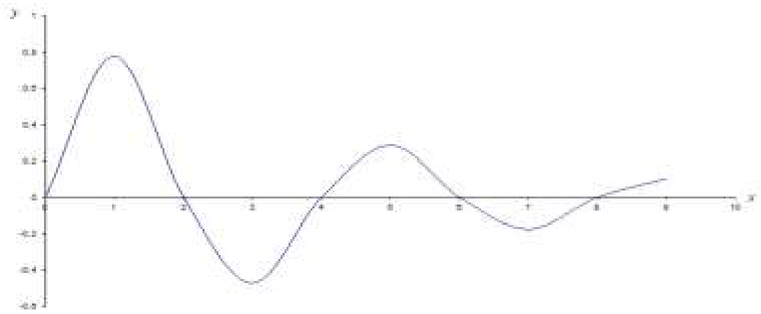

Figure 2: *The transition process of the object*

Object output is divided into seven classes:

$$y \in \underbrace{[-0.47,-0.3]}_{d1} \cup \underbrace{[-0.30,\ -0.05]}_{d2} \cup \underbrace{[-0.05,\ 0.15]}_{d3} \cup \underbrace{[0.15,0.30]}_{d4} \cup$$
$$\cup \underbrace{[0.3\ \ 0.45]}_{d5} \cup \underbrace{[0.45,\ 0.65]}_{d6} \cup \underbrace{[0.65,0.78]}_{d7}$$

The studied problem consists of synthesizing the object described in (1) according to 5 rules. Let us assume that the weights of the rules are equal to 0 or 1.

After linguistic interpretation, this rule will be described as follows:

IF x is close to 2.8 then $y \in d_1$

IF x is close to 6.9 then $y \in d_2$

IF x is close to 0 or x is close to 8.8 or x is close to 10 then $y \in d_3$

IF x is close to 5 then $y \in d_4$
IF x is close to 0.9 then $y \in d_7$

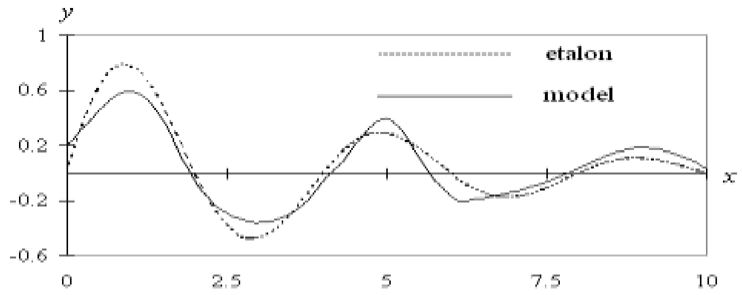

Figure 3. *Comparison of linguistic model and benchmark*

IV. Conclusıon

The approximation algorithm for objects with discrete output is designed in the following order: specify the vector of values of input data, enter the membership function of fuzzy terms used in the fuzzy knowledge base and determine the values of this function for the given values of input data; calculate the multivariate membership function for all solution values of the output variable using logic equations.

The proposed algorithm uses the idea of identifying linguistic terms by the maximum of the membership function. This idea is generalized across the entire knowledge matrix.

Research of the fuzzy system based on experimental data has shown that it is more effective than traditional methods.

CONFLICT OF INTEREST.

Authors declare that they do not have any conflict of interest.

References

[1] Zadeh Lotfi A. "Fuzzy logic, neural networks, and soft computing". *Communications of the ACM*. **37** (3): 1994, 77–84. doi: 10.1145/175247.175255. ISSN 0001-0782.

[2] Lütfi Z., Aliyev R.A. “Fuzzy Logic Theory and Applications”, World Scientific, 2018

[3] Aliev R.A., Aliev R.R. Soft Computing. Bakı, Chashıoglu, 2004, 484 s.

[4] López V.et al. A hierarchical genetic fuzzy system based on genetic programming for addressing classification with highly imbalanced and borderline data-sets / Knowledge-Based Systems 38 (2013) pp.85–104.

[5] Gardashova L.A., Mammadova K.A. Optimal Implicatıon Based Fuzzy Control System for a Steam Generator. 15th International conference on applications of fuzzy systems, Soft computing and artificial intelligence tools. ICAFS 2022. August 26 - 27 2022, Budva – Montenegro.

[6] Bo Zhou, Anna Konstorum et al. A Hierarchical Modeling Approach to Data Analysis and Study Design in a Multi-Site Experimental FMRI Study. Psychometrika. 2013 Apr; 78(2): 260–278. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4142354

[7] Medsker, Larry R. (2012-12-06). *Hybrid Intelligent Systems*. Springer Science & Business Media. ISBN 978-1-4615-2353-6.

[8] Van den Noortgate, W., & Onghena, P. (2003). Combining single-case experimental data using hierarchical linear models. *School Psychology Quarterly*, *18*(3), 325–346. https://doi.org/10.1521/scpq.18.3.325.2257

[9] İhsan Kaya, Murat Çolak, Fulya Terzi. A comprehensive review of fuzzy multi criteria decision making methodologies for energy policy making. Energy Strategy Reviews. Energy Strategy Reviews 24 (2019), pp.207–228. https://www.sciencedirect.com/science/article/pii/S2211467X19300252

[10] Sugeno, M., Kang, T. And Kang, G. Structure Identification of Fuzzy Model. Fuzzy Sets and Systems, 2017, 28, 15-33.

[11] An application of TOPSIS for ranking internet web browsers. Shahram Rostampour. Department of Management Science, Islamic Azad University, Central Branch, Tehran, Iran. 2012, p. 7