Data-driven Fuzzy Modeling Method and its Applications

Outline

- Brief resumes
- Fuzzy set theory
- Some acquired results
- A practical project

Degang Wang

- Received the PhD degree in Applied Mathematics from Beijing Normal University, China, in 2008.

- Since 2013, has been an associate professor in the School of Electronic and Information Engineering, Dalian University of Technology.

Address: No.2 Linggong Road, Ganjingzi District, Dalian City, Liaoning Province, P.R.C., 116024
E-mail: wangdg@dlut.edu.cn

Dalian University of Technology (DUT) is a national key university administrated directly under the Ministry of Education of China, which is also sponsored by Project 211 and Project 985.

DUT has now 4 national key primary disciplines (Dynamics, Hydraulic engineering, Chemical Engineering and Technology, Management Science and Engineering, covering 15 secondary disciplines), and 6 national key secondary disciplines (Computational Mathematics, Plasma Physics, Mechanical Manufacturing and Automation, Structural Engineering, Structural Analysis for Industrial Equipment, Environmental Engineering).

http://en.dlut.edu.cn/

Wenyan Song

- Received the PhD degree in Applied Mathematics from Beijing Normal University, China, in 2007.

- Since 2011, has been an associate professor in the School of Economics, Dongbei University of Finance and Economics.

Address: No.217 Jianshan Street, Shahekou District, Dalian City, Liaoning Province, P.R.C., 116025
E-mail: agathasong@126.com

Dongbei University of Finance and Economics (DUFE) is one of main institutes of finance and economy in China.

Applied economics, business administration, statistics and public management are dominant disciplines of DUFE.

Fuzzy Set and Fuzzy Logic

Berkeley | EECS
Electrical Engineering and Computer Sciences

Home / People / Faculty / Lotfi A. Zadeh

Lotfi A. Zadeh

Professor Emeritus

Research Areas

- Artificial Intelligence (AI)
- Control, Intelligent Systems, and Robotics (CIR)
- soft computing, knowledge-based systems

Research Centers

- Berkeley Artificial Intelligence Research Lab (BAIR)

- L. A. Zadeh, "Fuzzy sets," Information and Control, vol. 8, no. 3, pp. 338-353, June 1965.

• Fuzzy sets as generic constructs – building conceptual blocks using which we describe systems (develop models) in a meaningful way.

Each fuzzy set comes with a well-delineated semantics (meaning), e.g., *small* – *medium* – *large* error.

- Fuzzy logic is an approach to computing based on "degrees of truth" rather than the usual "true or false" (1 or 0) Boolean logic on which the modern computer is based.

Fuzzy logic works with membership values in a way that mimics Boolean logic.

If-Then rules map input or computed truth values to desired output truth values.

Fuzzy rule-based models

Generic format of fuzzy rule:

If input variable is *A* **then** output variable is *B*.

- *A* and *B*: descriptors of pieces of knowledge
- rule: expresses a relationship between inputs and outputs

Example:

If the temperature is high **then** the electricity demand is high.

If the income is low **then** the tax is low.

A collection of "if-then" (conditional) statements in the form:

If input variable is A_i **then** output variable is B_i.

(condition) (conclusion)

where A_i, B_i are fuzzy sets and $i=1, 2, \ldots, n$.

Construction of fuzzy rule-based models

Main steps:

- Specification of the fuzzy variables to be used and their quantification.
- Association of fuzzy variables using fuzzy sets.
- Formalization of rules using fuzzy relations and their aggregation.
- Construction of mapping (inference, approximate reasoning) on a basis of existing facts.

https://se.mathworks.com/products/fuzzy-logic.html

Contact Us How to Buy Log In

MathWorks® Products Solutions Academia Support Community Events

Fuzzy Logic Toolbox

Search MathWorks.com

Overview | Features | Code Examples | Videos | Webinars | What's New | Product Pricing | Trial software | Contact sales

Design and simulate fuzzy logic systems

Fuzzy Logic Toolbox™ provides functions, apps, and a Simulink® block for analyzing, designing, and simulating systems based on fuzzy logic. The product guides you through the steps of designing fuzzy inference systems. Functions are provided for many common methods, including fuzzy clustering and adaptive neurofuzzy learning.

The toolbox lets you model complex system behaviors using simple logic rules, and then implement these rules in a fuzzy inference system. You can use it as a stand-alone fuzzy inference engine. Alternatively, you can use fuzzy inference blocks in Simulink and simulate the fuzzy systems within a comprehensive model of the entire dynamic system.

R2017b

» Explore the latest features for this product

Free eBook: Machine Learning with MATLAB

» Download now.

Key Features

- Fuzzy Logic Design app for building fuzzy inference systems and viewing and analyzing results
- Membership functions for creating fuzzy inference systems
- Support for AND, OR, and NOT logic in user-defined rules
- Standard Mamdani and Sugeno-type fuzzy inference systems
- Automated membership function shaping through neuroadaptive and fuzzy clustering learning techniques
- Ability to embed a fuzzy inference system in a Simulink model
- Ability to generate embeddable C code or stand-alone executable fuzzy inference engines

1. Approximation to a class of non-autonomous systems by dynamic fuzzy inference marginal linearization method

Let us consider the following non-autonomous system.

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t), \\ \mathbf{x}(t_0) = \mathbf{x}_0. \end{cases} \tag{14}$$

where $\mathbf{x}_0 = (x_1(t_0), \ldots, x_n(t_0))$, $\mathbf{f} = (f_1, \ldots, f_n)$. $\mathbf{x} = (x_1, \ldots, x_n)^T$ is the state vector and x_i $(i = 1, \ldots, n)$ are state variables. X_1 - $\times \cdots \times X_n$ and $\dot{X}_1 \times \cdots \times \dot{X}_n$ are the universes of $\mathbf{x}$ and $\dot{\mathbf{x}}$ respectively. Similarly, we also assume that X_i $(i = 1, \ldots, n)$ and U are real number intervals respectively, i.e., $X_i = [a_i, b_i]$, $i = 1, \ldots, n$, $U = [t_0, t_b]$.

$$\begin{cases} \text{If } (x_1, \ldots, x_n, t) \text{ is } A_{1j_1} \times \cdots \times A_{nj_n} \times T_k \text{ then } \dot{x}_1 \text{ is } B_{k1j_1j_2\cdots j_n}; \\ \text{If } (x_1, \ldots, x_n, t) \text{ is } A_{1j_1} \times \cdots \times A_{nj_n} \times T_k \text{then } \dot{x}_2 \text{ is } B_{k2j_1j_2\cdots j_n}; \\ \vdots \\ \text{If } (x_1, \ldots, x_n, t) \text{ is } A_{1j_1} \times \cdots \times A_{nj_n} \times T_k \text{then } \dot{x}_n \text{ is } B_{knj_1j_2\cdots j_n}. \end{cases}$$

Theorem 3. *Suppose fuzzy sets $A_{ij_i}(x_i)$ $(i = 1, \ldots, n;\ j_i = 1, \ldots, p_i)$ and fuzzy sets $T_k(t)(k = 1, \ldots, m)$ are respectively chosen as triangle-shaped membership functions, then the time-variant fuzzy system determined by rules (15) and DFIML method can be expressed as follows:*

$$\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}, t), \tag{16}$$

where $g_j(\mathbf{x},t) = a_{j0}t + a_{j1}x_1 + \cdots + a_{jn}x_n + b_{j1}tx_1 + \cdots + b_{jn}tx_n + c_j$:

Theorem 6. *Consider the Eq. (14) and assume that $\mathbf{f}(\mathbf{x}, t)$ is piecewise continuous in t and locally Lipschitz in $\mathbf{x}$ on $C \times [t_0, +\infty)$, where $C = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}_0\| < d\}$. Let W be a compact subset of C and $\mathbf{x}_0 \in W$. Suppose that for all $t \in [t_0, +\infty)$, the solution of Eq. (14) lies in W. Then, given $\varepsilon > 0$ and $t \in [t_0, +\infty)$, there exists a time-variant fuzzy system generalized by DFIML method satisfies*

$$\|\phi(t') - \psi(t')\| < \varepsilon, \quad \forall t' \in [t_0, t]$$

where $\psi(t')$ is the solution of time variant fuzzy system.

Theorem 7. *Consider the Eq. (14) and assume that $\mathbf{f}(\mathbf{x}, t)$ is piecewise continuous in t and locally Lipschitz in $\mathbf{x}$ on $C \times [t_0, +\infty)$, where $C = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x} - \mathbf{x}_0\| < d\}$. Let W be a compact subset of C and $\mathbf{x}_0 \in W$. Suppose that for all $t \in [t_0, +\infty)$, the solution of Eq. (14) lies in W and for any $\mathbf{x} \in W$, $\int_{t_0}^{+\infty} \mathbf{f}(\mathbf{x}, t)dt < \infty$. Then, given $\varepsilon > 0$, there exists a time-variant fuzzy system generalized by DFIML method satisfies*

$$\|\phi(t) - \psi(t)\| < \varepsilon, \quad \forall t \in [t_0, +\infty)$$

where $\psi(t)$ is the solution of time variant fuzzy system.

Fuzzy set and Artificial Intelligence

- **Computational intelligence**

Neural networks are an example of soft computing.

Other soft computing approaches to AI include fuzzy systems, evolutionary computation and many statistical learning methods.

2. Fuzzy Wavelet Neural Network

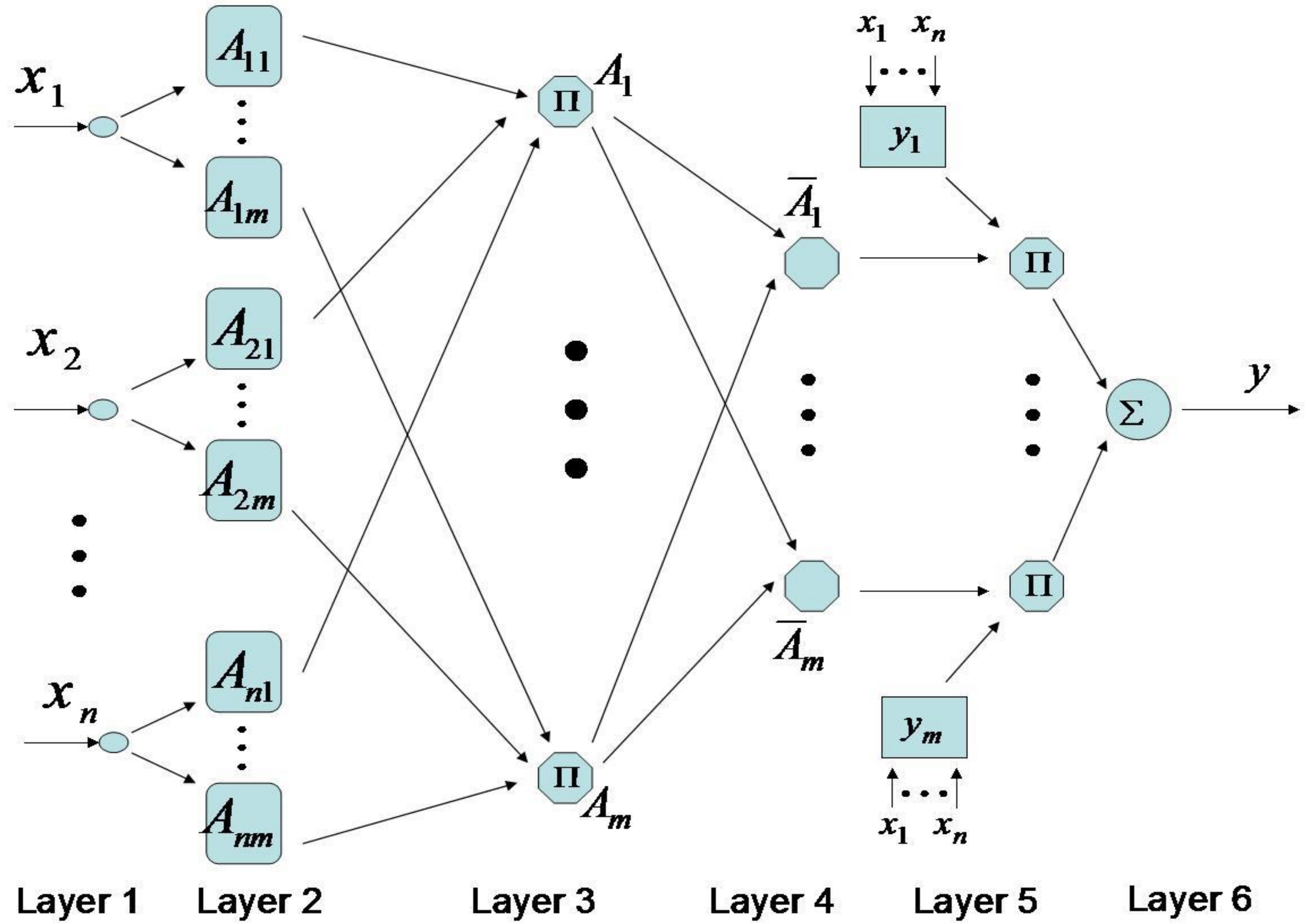
$$Rule\ j: if\ x_1 \text{ is } A_{1j}\ and\ x_2\ is\ A_{2j} \cdots and\ x_m\ is\ A_{mj}$$

$$then\ y_j = \sum_{i=1}^{n} w_{ij}\varphi_{ij}(x_i)$$

A_{ij} -- membership function

$\varphi_{ij}(x_i)$ -- wavelet function

y_j -- consequent of the *j*-th rule



x_1
x_2
x_n
A_{11}
A_{1m}
A_{21}
A_{2m}
A_{n1}
A_{nm}
Π
A_1
A_m
$\bar{A}_1$
$\bar{A}_m$
x_1 x_n
y_1
y_m
Σ
y
Layer 1
Layer 2
Layer 3
Layer 4
Layer 5
Layer 6

Mathematical model of FWNN

$$y = \sum_{j=1}^{m} \bar{A}_j(\mathrm{x}) \cdot \left(\sum_{i=1}^{n} w_{ij} \varphi_{ij}(x_i) \right)$$

$$= \sum_{j=1}^{m} \bar{A}_j(\mathrm{x}) \cdot \left(\sum_{i=1}^{n} w_{ij} \cdot \left| a_{ij} \right|^{-\frac{1}{2}} \cdot \left(1 - \left(\frac{x_i - b_{ij}}{a_{ij}} \right)^2 \right) \cdot \exp\left(-\frac{1}{2} \cdot \left(\frac{x_i - b_{ij}}{a_{ij}} \right)^2 \right) \right)$$

$$\bar{A}_i = A_i \bigg/ \sum_{i=1}^{n} A_i \quad \longleftarrow \quad A_j(\mathrm{x}) = \prod_{i=1}^{n} A_{ij}(x_i) \quad \longleftarrow \quad A_{ij}(x_i) = \exp\left(-\frac{(x_i - c_{ij})^2}{\sigma_{ij}^2}\right)$$

$$\varphi_{ij}(x_i) = \left| a_{ij} \right|^{-\frac{1}{2}} \cdot (1 - z_{ij}^2) \cdot \exp\left(-\frac{z_{ij}^2}{2}\right) \qquad z_{ij} = \frac{x_i - b_{ij}}{a_{ij}}$$

Parameter learning method for FWNN

From the structure of FWNN, we have to identify these parameters:

- linear parameters

$$y_l = \sum_{i=1}^{n}\sum_{j=1}^{m} \bar{A}_j(x_l)\cdot\varphi_{ij}(x_i^l)\cdot w_{ij} \longrightarrow w_{ij}$$

generalized least square

- nonlinear parameters

$$A_{ij}(x_i) = \exp\left(-\frac{(x_i - c_{ij})^2}{\sigma_{ij}^2}\right) \longrightarrow c_{ij}\ \ \sigma_{ij}$$

$$\varphi_{ij}(x_i) = \left|a_{ij}\right|^{-\frac{1}{2}}\cdot(1 - z_{ij}^2)\cdot\exp\left(-\frac{z_{ij}^2}{2}\right) \longrightarrow a_{ij}\ \ b_{ij}$$

$$z_{ij} = \frac{x_i - b_{ij}}{a_{ij}}$$

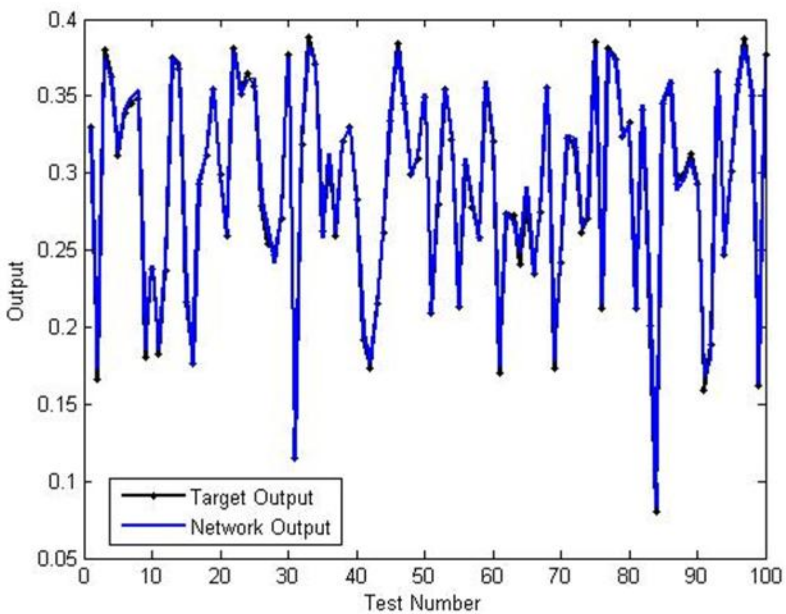
gradient descent

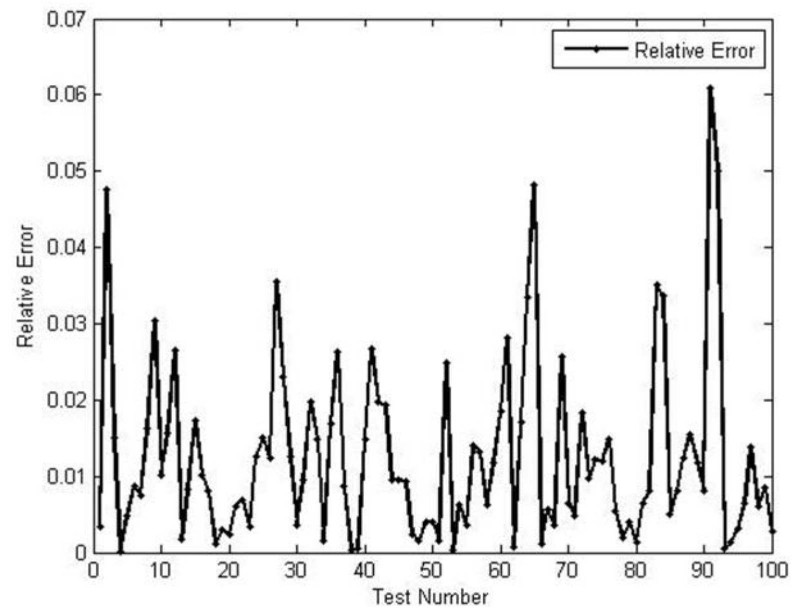
Example. Nonlinear function

Consider the following nonlinear function

$$y = \sqrt{64 - 81((x_1 - 0.6)^2 + (x_2 - 0.5)^2)} \Big/ 9 - 0.5 \qquad x_1, x_2 \in [0,1]$$

In the simulation, 1000 samples are randomly generated in $[0,1]\times[0,1]$, 3 fuzzy rules are proposed to construct FWNN.



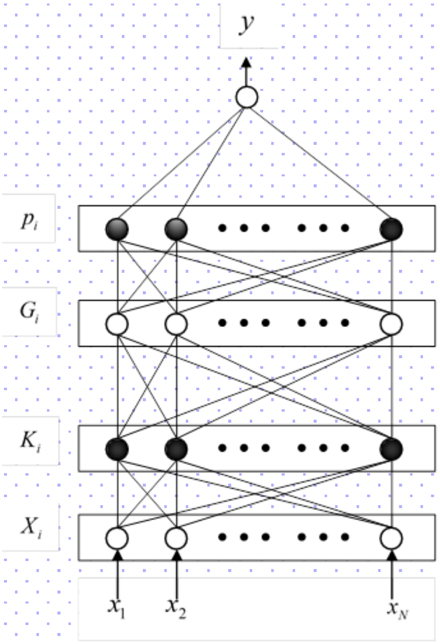


3. Kernel Logistic Neural Network

- The first layer of the KLNN-RBM model is the input layer.
- The second layer is the transformation layer.
  Kernel function is added in this layer. Since the dimensions of the input parameters will increase , Principal Component Analysis (PCA) is carried out on the dimension reduction of the kernel function $K(x_i, x_j) \subset \mathbb{R}^{N\times N}$.
- The third layer is the BP neural network layer.
  The input is $k_i = [k_{i1} \quad \cdots \quad k_{in}]$, which is obtained from PCA.
  The output of this layer is represented by:

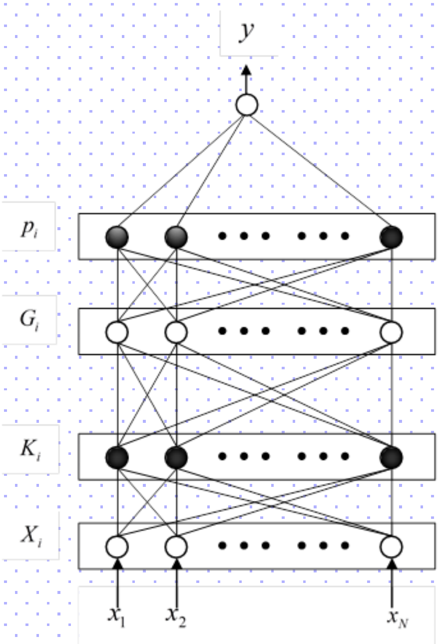
$$G_i = \sigma(\theta + \omega^T{}_h \cdot k_i)$$

where $\sigma(u) = 1/(1+\exp(-u))$, ω_h and θ is the weight and threshold of the network.



Contrastive Divergence

- The fourth layer is the logistic regression layer. And the input of the logistic function is the composite function model which is the output of the third layer. For multi-class classification, in this paper, we use the one-versus-all(OVA) method to handle it. So logistic function can be written by:



$$p_m(y=1)=\frac{1}{E}\exp(\beta_m^T\cdot G_i+\alpha_m)$$

$$p_M(y=M)=1-(p_1+\cdots+p_{M-1})$$

where $E=1+\sum_{m=1}^{M-1}\exp\left(\beta_m^T\cdot G_i+\alpha_m\right)$, β_m and α_m is the weights and thresholds of the logistic function, m=1, 2$\cdots M$-1 , M is the number of classes.

- The fifth layer is the output layer.

stochastic gradient descent

$$y=\arg\max_{c\in\{1,2,\cdots,M\}}p_c$$

Example.

Datasets: University of California Irvine (UCI) (http://archive.ics.uci.edu/ml/index.php)

Description of UCI datasets

Data	Classes	Instances	Features
German	2	1000	24
PID	2	768	8
Heart	2	270	13
Wine	3	178	13
Iris	3	140	4
Derm	6	358	34

- Lichman, M. (2013). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science.

- The linear kernel and the Gaussian RBF kernel are respectively used to construct KLNN-RBM model, i.e.

$$K_1(x_i, x_j) = x_i \cdot x_j^T \qquad K_2(x_i, x_j) = \exp(-\frac{1}{2\sigma^2}\|x_i - x_j\|)^2$$

- For the binary UCI datasets, we use 5-fold cross-validation. The number of hidden units are taken as $H=3$. The number of iterations for RBM and stochastic gradient descent method are taken as 100.

Simulation results of binary UCI datasets

Data	Gaussian KLNN -RBM	Linear KLNN -RBM	RBF -R [14]	RBF -N [14]	RBF -WTA [14]	DKP [15]	ELM [16]
German	77.00	76.70	69.1	69.9	66.3	74.1	75.4
PID	77.87	76.96	75.3	72.1	73.8	74.7	70.9
Heart	84.81	84.07	81.9	80.5	80.6	79.9	74.7

- The linear kernel and the Gaussian RBF kernel are respectively used to construct KLNN-RBM model, i.e.

$$K_1(x_i, x_j) = x_i \cdot x_j^{\ T} \qquad K_2(x_i, x_j) = \exp\left(-\frac{1}{2\sigma^2}\left\|x_i - x_j\right\|\right)^2$$

- For the multi-class UCI datasets, we use 10-fold cross-validation. The number of hidden units are taken as $H=4$. The number of iterations for RBM and stochastic gradient descent method are taken as 100.

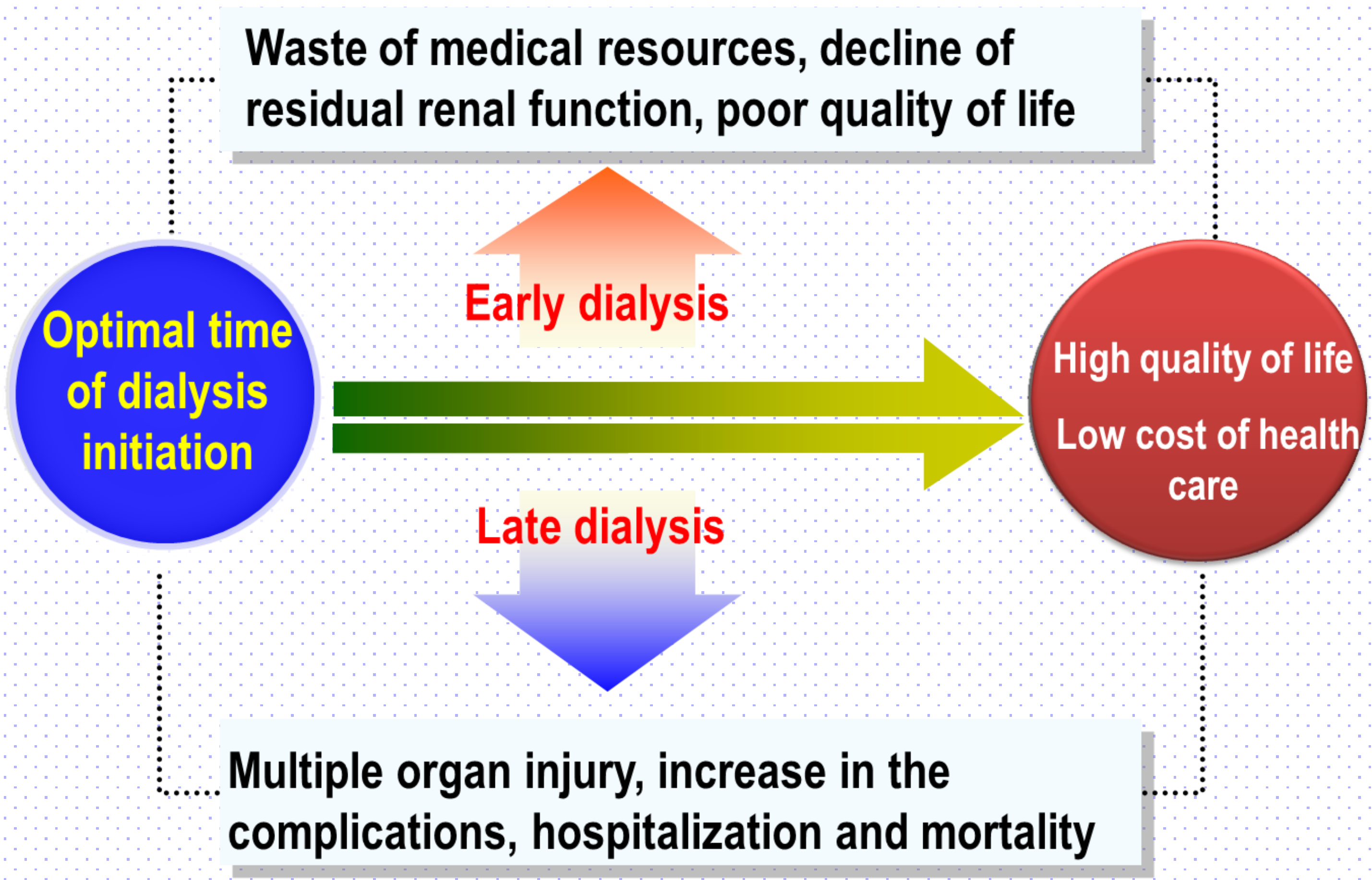
Simulation results of multi-class UCI datasets

Data	Gaussian KLNN -RBM	Linear KLNN -RBM	TR-KLR[6]	Non-linear-MLETSVM [17]	Linear-MLETSVM [17]
Wine	100	100	99.47	100	100
Iris	99.33	100	98.00	97.75	96.35
Derm	99.17	99.71	97.45	94.71	93.85

Application Case

Timing of dialysis initiation

--key factor of prognosis and medical cost



Data collection

Demographic clinical and laboratory data at the initiation of dialysis

Outcome----status within 3-yr of initiation (survival time or 36m)

Statistical analysis

Mann-Whitney U test

χ^2 test

ROC curve

Development of the CHDSE equation

Fuzzy system theory

variables

KLNN-RBM

Basic variables—

Age, sex, BUN, Scr, Alb

Equation development

Sensitivity

Specificity

Diagnostic accuracy

Equation evaluation

Compare the CHDSE and MDRD equation using ROC

Thank You.