Contemporary Data Processing Technology (CCOD)
Lab 6 (October 13, 2016)
Savchuk Artem, AS-36

We have Iris data base (Table - 1). In this table represented 3 class of iris flower with 4 parameters(x1, x2, x3 and x4). Where x1, for example, is sepal-length, x2 is sepal-width and etc. Our first goal is to analyze the table and create 3 membership functions of small, medium and large. For this we should take all values of table and put them into one group. This group will consist of 3 x 4 x 17 values. Then we should divide this big group into 3 another groups of small, medium and large. Then we should define Gaussian membership function for each group.
Gaussian membership function:

$$f(x) = \exp\left\{-\frac{(x-avg)^2}{std}\right\};$$

Where avg is the average of the group, std is standard deviation of the group, N is number of values in group.

$$avg = \frac{\sum_{i=1}^{N} x_i}{N};$$

$$std = \sum_{i=1}^{N} \frac{(x_i - avg)^2}{N};$$

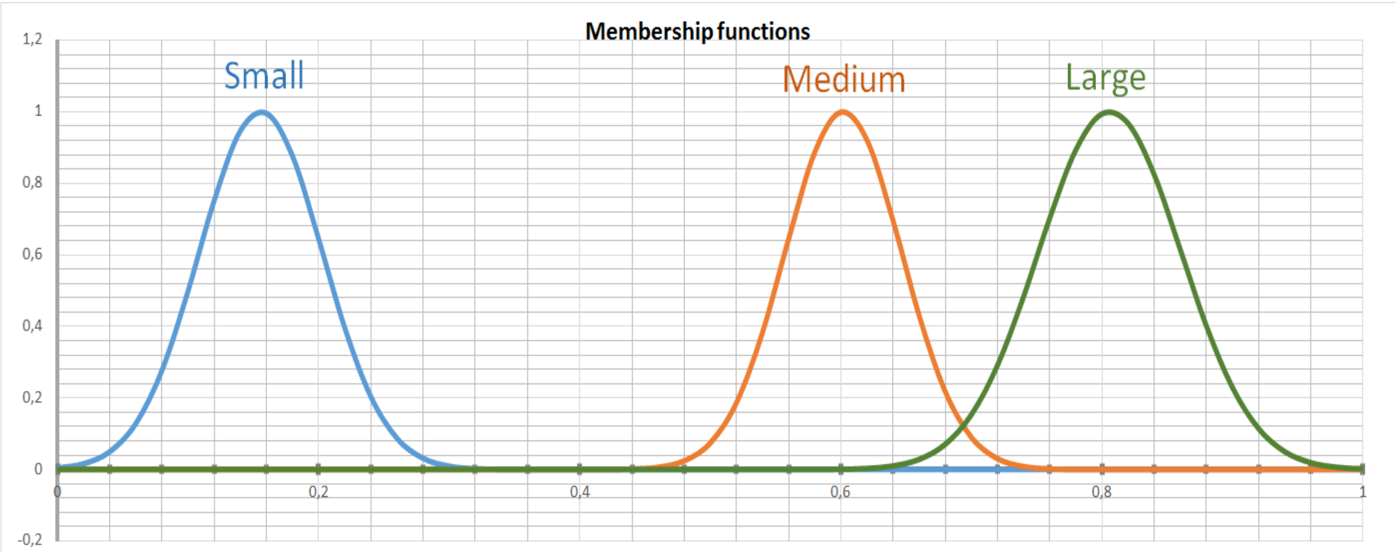
Setosa				Versicolor				Virginica			
x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4
0.56	0.66	0.20	0.08	0.84	0.66	0.67	0.52	0.85	0.57	0.84	0.72
0.62	0.70	0.22	0.04	0.66	0.61	0.57	0.56	0.91	0.82	0.88	1.00
0.68	0.84	0.22	0.08	0.63	0.45	0.51	0.40	0.82	0.73	0.74	0.80
0.61	0.77	0.23	0.08	0.75	0.68	0.61	0.60	0.81	0.61	0.77	0.76
0.61	0.68	0.20	0.04	0.76	0.50	0.58	0.40	0.86	0.68	0.80	0.84
0.54	0.68	0.16	0.04	0.77	0.66	0.68	0.56	0.72	0.57	0.72	0.80
0.73	0.91	0.17	0.08	0.71	0.66	0.52	0.52	0.73	0.64	0.74	0.96
0.72	1.00	0.22	0.16	0.85	0.70	0.64	0.56	0.81	0.73	0.77	0.92
0.68	0.89	0.19	0.16	0.71	0.68	0.65	0.60	0.82	0.68	0.80	0.72
0.65	0.80	0.20	0.12	0.73	0.61	0.59	0.40	0.97	0.86	0.97	0.88
0.72	0.86	0.25	0.12	0.78	0.50	0.65	0.60	0.97	0.59	1.00	0.92
0.65	0.86	0.22	0.12	0.71	0.57	0.57	0.44	0.76	0.50	0.72	0.60
0.68	0.77	0.25	0.08	0.75	0.73	0.70	0.72	0.87	0.73	0.83	0.92
0.65	0.84	0.22	0.16	0.77	0.64	0.58	0.52	0.71	0.64	0.71	0.80
0.58	0.82	0.14	0.08	0.80	0.57	0.71	0.60	0.97	0.64	0.97	0.80
0.65	0.75	0.25	0.20	0.77	0.64	0.68	0.48	0.80	0.61	0.71	0.72
0.61	0.77	0.28	0.08	0.81	0.66	0.62	0.52	0.85	0.75	0.83	0.84

Table - 1

I calculated avg and std for 3 groups and got 3 membership functions of small, medium, and large:

- Gaussian membership function of small is $f(x) = \exp\left\{-\frac{(x-0.1556)^2}{0.00447}\right\}$;
- Gaussian membership function of medium is $f(x) = \exp\left\{-\frac{(x-0.6019)^2}{0.004}\right\}$;
- Gaussian membership function of large is $f(x) = \exp\left\{-\frac{(x-0.806)^2}{0.006}\right\}$;

Then I drew these Gaussian membership functions of small, medium and large:



Pic-1

Next goal is to create one rule for each class of iris flower. Using the data in Table-1 and membership function on Pic-1 I created 3 Takagi-Sugeno rules by my feeling:

1. IF X1=<u>medium</u> AND X2=<u>large</u> AND X3=<u>small</u> AND X4=<u>small</u> THEN y = 1
2. IF X1=<u>large</u> AND X2=<u>medium</u> AND X3=<u>medium</u> AND X4=<u>medium</u> THEN y = 2
3. IF X1=<u>medium</u> AND X2=<u>large</u> AND X3=<u>large</u> AND X4=<u>large</u> THEN y = 3

Then we should evaluate our 3 rules using algorithm Takagi-Sugeno and data in Table-2.

Setosa				Versicolor				Virginica			
x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4
0.65	0.80	0.20	0.08	0.89	0.73	0.68	0.56	0.80	0.75	0.87	1.00
0.62	0.68	0.20	0.08	0.81	0.73	0.65	0.60	0.73	0.61	0.74	0.76
0.59	0.73	0.19	0.08	0.87	0.70	0.71	0.60	0.90	0.68	0.86	0.84
0.58	0.70	0.22	0.08	0.70	0.52	0.58	0.52	0.80	0.66	0.81	0.72
0.63	0.82	0.20	0.08	0.82	0.64	0.67	0.60	0.82	0.68	0.84	0.88
0.68	0.89	0.25	0.16	0.72	0.64	0.65	0.52	0.96	0.68	0.96	0.84
0.58	0.77	0.20	0.12	0.80	0.75	0.68	0.64	0.62	0.57	0.65	0.68
0.63	0.77	0.22	0.08	0.62	0.55	0.48	0.40	0.92	0.66	0.91	0.72

Table-2

Takagi-Sugeno Fuzzy formula

Takagi-Sugeno Singleton Consequence looks like this:

$$R_k\text{: If } x_1 \text{ is } A_1^k, \text{ and } x_2 \text{ is } A_2^k \text{ and } \cdots \text{ and } x_N \text{ is } A_N^k \text{ then } y \text{ is } g^k.$$

Where x – input, k – index of rule, A – membership function, N – number of parameters.

Estimation of y for an input $\mathbf{x} = (x_1, x_2, \cdots, x_N)$

$$y_j = \frac{\sum_{k=1}^{H}(M_k(\mathbf{x}) \cdot g_k)}{\sum_{k=1}^{H} M_k(\mathbf{x})}$$

where

$$M_k(\mathbf{x}) = \prod_{i=1}^{N} \mu_{ik}(x_i)$$

where μ_{ik} is i-th attribute of k-th rule

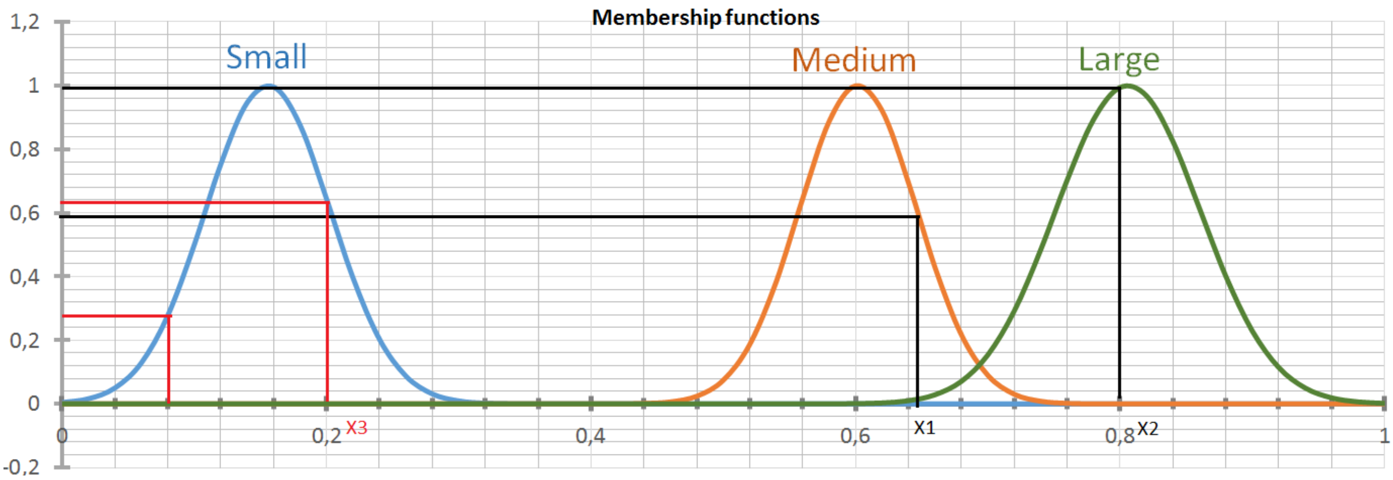
When we calculate y_j, we can determine class:

$$y = \begin{cases} 1 \ \ldots \text{ if } \hat{y} < 1.5 \\ 2 \ \ldots \text{ if } 1.5 <= \hat{y} < 2.5 \\ 3 \ \ldots \text{ if } 2.5 <= \hat{y} \end{cases}$$

Where 1, 2 and 3 is number of class.

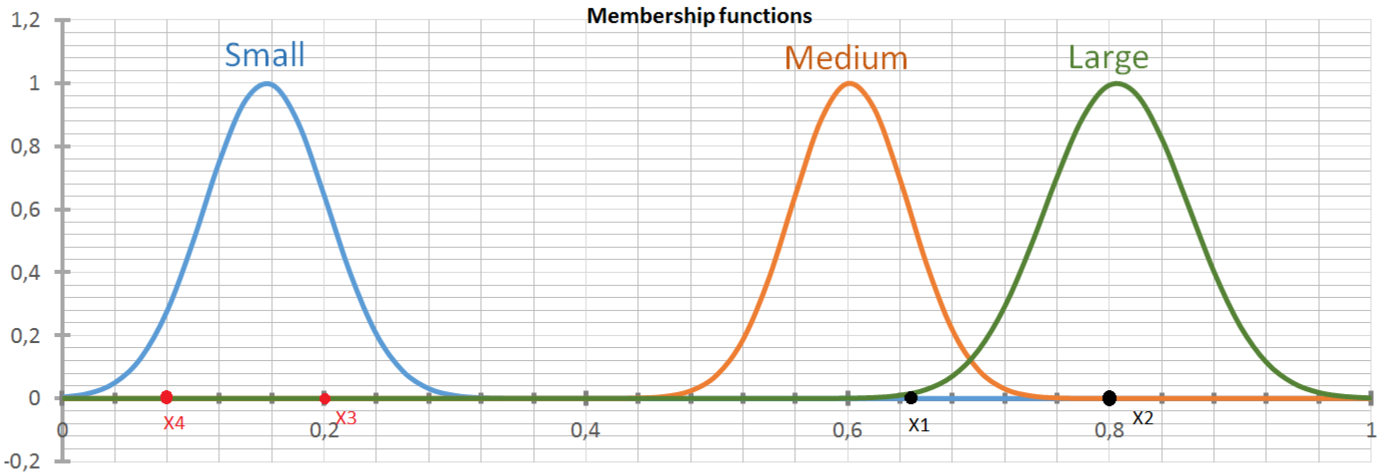
For example, we consider a flower with a set of the following: x1 = 0.65, x2 = 0.8, x3 = 0.2 and x4 = 0.08; Then we are going to implement algorithm Takagi-Sugeno and determine which class.

Rule 1:



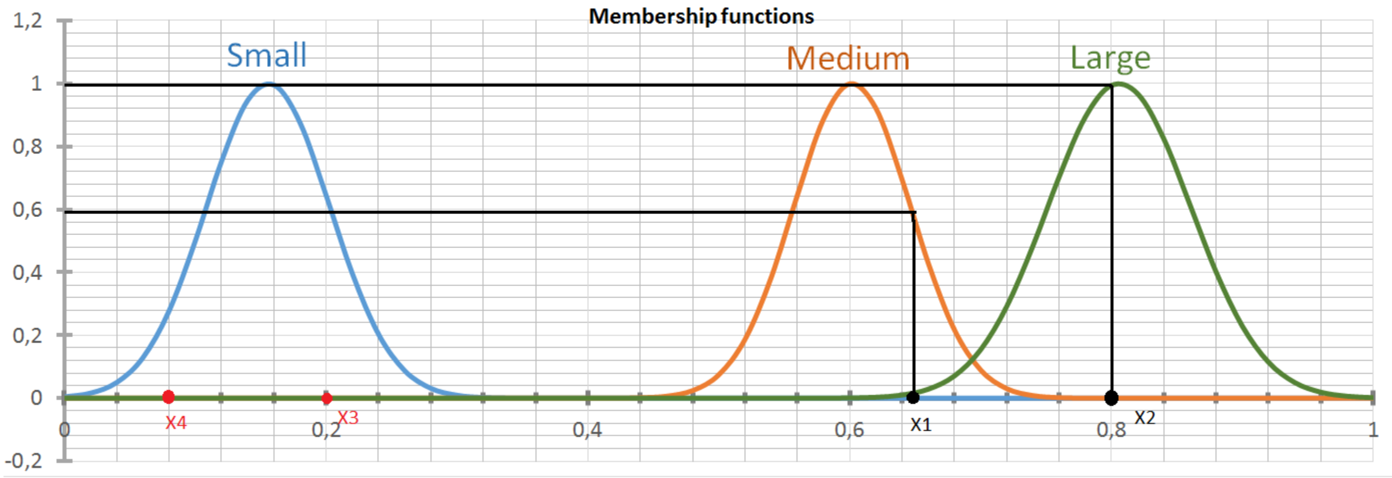
$\mu_{medium}(x1) = \mu_{medium}(0.65) = 0.59;$
$\mu_{large}(x2) = \mu_{large}(0.8) = 0.99;$
$\mu_{small}(x3) = \mu_{small}(0.2) = 0.63;$
$\mu_{small}(x4) = \mu_{small}(0.08) = 0.28;$
$M_1 = \mu_{medium}(x1) * \mu_{medium}(x2) * \mu_{small}(x3) * \mu_{small}(x4) = 0.59 * 0.99 * 0.63 * 0.28 = \mathbf{0.103};$

Rule 2:



$\mu_{large}(x1) = 0.017;$
$\mu_{medium}(x2) = 0.005;$
$\mu_{medium}(x3) = 2.9 * 10^{-8};$
$\mu_{medium}(x4) = 2.6 * 10^{-30};$
$M_2 = \mu_{medium}(x1) * \mu_{medium}(x2) * \mu_{small}(x3) * \mu_{small}(x4) = 0.017 * 0.005 * 2.9 * 10^{-8} * 2.6 * 10^{-30} = 5.083 * 10^{-42} \approx \mathbf{0};$

Rule 3:



$\mu_{medium}(x1) = 0.59;$
$\mu_{large}(x2) = 0.99;$
$\mu_{large}(x3) = 2.67 * 10^{-27};$
$\mu_{large}(x4) = 7.06 * 10^{-39};$
$M_3 = \mu_{medium}(x1) * \mu_{medium}(x2) * \mu_{small}(x3) * \mu_{small}(x4) = 0.59 * 0.99 * 2.67 * 10^{-27} * 7.06 * 10^{-39} = 1.1 * 10^{-65} \approx \mathbf{0};$

$$y = \frac{M_1 * 1 + M_2 * 2 + M_3 * 3}{M_1 + M_2 + M_3} = \frac{0.103 + 0 * 2 + 0 * 3}{0.103 + 0 + 0} = 1$$

As y < 1.5, the flower with set of parameters {x1 = 0.65, x2 = 0.8, x3 = 0.2, x4 = 0.08} belongs to class 1.

After I did it with all the values of the Table-2, according to the result I created table:

No.	Family A	Family B	Family C	Evaluation
#1	A	B	C	Good
#2	A	B	C	Good
#3	A	B	C	Good
#4	A	B	C	Good
#5	A	B	C	Good
#6	A	B	C	Good
#7	A	B	B	Not Good
#8	A	B	B	Not Good
Success rate	100,00%	100,00%	75,00%	75,00%

Table-3