***Machine Learning-*** *Homework 1*

Deadline: 2021/03/29

1. **Gradient descent**: A fuzzy approximator is constructed by a set of fuzzy rules as



Generally, *θl* is a fuzzy singleton (TS fuzzy model). Another type is to use a linear combination of input variables. In that case, usually, the recursive least square (RLS) approach (or recursive Kalman filter) can be used to identify those coefficients. The fuzzy systems with the center-of area like defuzzification and product inference can be obtained as

.

If the fuzzy system is utilized to approximate a given data set or training pattern (**x**, *y*). *Please derive the parameters update law by gradient descent method for θl* .

1. Consider the given data set with 119 inputs and one output.

(a) Please use the multiple linear regression to solve the regression of dataset.

(b) Please calculate the simple linear regression for each input variable and output. And then rank them according training root mean square error.

(c) Pearson's correlation coefficient is the test statistics that measures the statistical relationship, or association, between two continuous variables. It is known as the best method of measuring the association between variables of interest because it is based on the method of covariance. Please calculate the Pearson's correlation coefficient and rank according the values of correlation coefficients.

(d) Compare the results in ranking of part (b) and (c), what do you observe?

(e) According literature [R1], the authors conclude that we can select the input variables for regression is ranking of RMSE, this means that the influence of variables can be determined by regression RMSE. Please find the suitable input variables to minimize the regression RMSE.

[R1] K. Mohammadi, S. Shamshirband, D. Petković, P. L. Yee, and Z. Mansor, "Using ANFIS for selection of more relevant parameters to predict dew point temperature," *Applied Thermal Engineering,* vol. 96, pp. 311-319, 2016.