Intelligent systems
1º Semester 2023/2024

Assignment 1

Camilo Andres Plata Uribe – 108565
Flavio Lussueki Lazaro – 107589

1. What is the difference between a characteristic function and a membership function?

A characteristic function maps elements from a set to binary values, usually the set {0,1}. It indicates whether a particular element belongs to a set or not in classical logic. In contrast, a membership function maps each element of the universe of discourse to a membership grade in the interval [0,1]. It assigns degrees of membership to elements in a fuzzy set.

2. Consider the two fuzzy sets in the Universe of Discourse X = {−8, −6, −4, −2, 0, 2, 4, 6, 8}:

$$\mu_A(x) = \frac{1}{1+|x|} \quad \textbf{and} \quad \mu_B = 1 - \frac{|x|}{20}$$

a) Are the membership functions valid in the given Universe?

For the given universe of discourse X, $\mu_A \in [1/9\,, 1]$ and $\mu_A \in [0.6, 1]$, so both fuzzy sets are valid.

b) Compute the α-cuts of A and B for $\alpha = 0.3$.

$A_{\alpha|\alpha=0.3} = \{x \in X \mid \mu_A(x) \geq 0.3\} = \{-2, 0, 2\}$

$B_{\alpha|\alpha=0.3} = \{x \in X \mid \mu_B(x) \geq 0.3\} = \{-8, -6, -4, -2, 0, 2, 4, 6, 8\} = X$

3. Classification Model for Wine dataset

Determination of first order Takagi-Sugeno classification model:

The given data set consist of 13 constituents found in each of the three types of wines. Thus, the classification problem includes 13 input variables and 3 classes (0,1,2). In order to compare the measurements on different scales it is essential to normalize the input variables. In this case, a Data box normalization was conducted (data normalization in a min-max [0,1] scale).

Firstly, to develop the classification of the data, three different classification models are established in which each of the classes is compared with the entire data set, assigning the value 1 as class membership and 0 as non-membership. Each model was assigned 70% of data for training and 30% of data for testing.

Then, for the three models the training data must be cluster, and the antecedent and consequent parameters must be estimated. With this it is possible to build a first order Takagi-Sugeno model for each classification model.

Particularly, the Input/Output space partition is obtained using fuzzy C-means clustering. The antecedent parameters are estimated from the cluster centers and partition matrix establish in the clustering process. the membership functions parameters are fitted to a gaussian membership function. Subsequently, the optimal consequent parameters are established by local optimization of consequents via least-mean squares estimation.

Finally, the test set is used to compute the class probabilities predictions of the models. The model output is given by the weighted fuzzy-mean, and the predicted classification is determined by an argmax estimation of the three outputs of the models.

Results:

Classification performance metrics

- Accuracy: 0.981
- Cohen's kappa: 0.972
- Matthew's correlation coefficient: 0.972