

Assignment: a fuzzy expert system for book sales prediction

Introduction

For editorials, predicting the number of sales of new books is a key factor. They must adjust the number of printed copies to distribute in order to avoid unsold books but also to have enough copies for the demand. ***This assignment exercise can be done in pairs of students.***

To make this prediction, different variables can be taken into account. Looking at the two papers indicated below, you must design a fuzzy expert system to calculate the number of books to print depending on the values of four variables: author visibility, publisher prestige, previous sales, and publishing period of year. Use the information given in the following papers to define these variables, as you consider appropriate.

J.C. Martín, E. Golobardes, X. Vilasis, CAIT: A Predictive Tool for Supporting the Book Market Operation Using Social Networks, *Women in Artificial Intelligence* book, MDPI, 2022. <https://www.mdpi.com/2076-3417/12/1/366>

X. Wang, B. Yucesoy, O. Varol, T. Eliassi-Rad, A.L. Barabási, Success in books, predicting book sales before publication, *EPJ Data Science*, 8, 31, 2019. <https://epjdatascience.springeropen.com/articles/10.1140/epjds/s13688-019-0208-6>

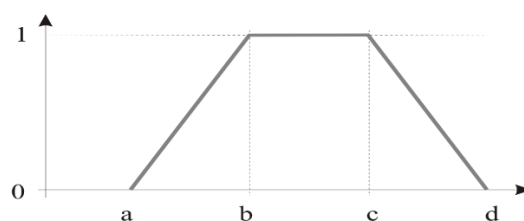
The tasks to be done are:

TASK 1: Definition of the input and output linguistic variables.

For each variable you must decide the reference scale, the number of terms and the corresponding fuzzy sets (by defining their membership functions). Remember they must satisfy the property of Fuzzy Partition.

You should use the information of the papers given above in order to support your design of the variables. Other sources of data can be used if you want.

You must use the trapezoidal functions to define each of the linguistic terms of the variables, using the (a,b,c,d) format, being its graphical representation:



TASK 2: Define the rules for the expert system

You must define a set of conjunctive rules. Decide appropriate premises and assign a degree of support to each rule. The rules must cover all possible combinations of input values, but you should use rules of different lengths. Support your definition of rules with evidences you find in the papers. Avoid inconsistencies in the rules. The number of rules should not exceed 30-35, in order make possible its manual analysis.

TASK 3: Implement the fuzzy expert system using Matlab.

Consider a Mamdani system (min as t-norm and max as t-conorm) and a Center of Area as defuzzification method. Validate the system using the 3D plot of the rules.

TASK 4: Define four different test cases and execute the system.

Explain the situation represented in each of the cases. You must configure appropriate test cases that represent different situations. Some of them must activate more than one label of the same variable. Report the results of each case with screenshots and explanations that justify the output obtained. Compare the results of the cases.

TASK 5: Answer the following questions and give a justification (no additional implementation is required):

1. Looking at your fuzzy expert system, what is the influence of the publishing period on the number of books if the author has great visibility?
2. What changes should you make to the system and which are the implications in the system if you want to consider also the age of the readers to whom the book is addressed?
3. If the author visibility depends on many criteria, how can we model it in this expert system?

Submission and deadline:

- A detailed report of the work done in the 5 previous tasks.
- Deliver the Matlab files of your systems too.
- Deadline: **12/12/2022**

Evaluation guide:

- TASK 1: definition of variables (correctness, motivation and justification) => 20 points
- TASK 2: definition of rules (correctness, motivation and justification) => 20 points
- TASK 3: implementation in Matlab => 20 points
- TASK 4: test cases, results and discussion => 20 points
- TASK 5: answers to final questions => 20 points