Assignment 3: a fuzzy expert system to detect phishing in websites

This assignment exercise can be done in pairs of students.

Detection of phishing websites is a critical safety measure for most online platforms. Phishing is a form of fraud in which the attacker tries to learn sensitive information of a person by imitating a known and reputed website. In the website, the user is requested to click malicious links or to fill in some fields with the purpose of obtaining data as personal identifier number, login credentials, and account information, among others. The use of Artificial Intelligence techniques to prevent phishing cyber-attacks is widespread (see [1][2][3]).

At the UCI repository (University of California Irvine – UCI) there is a public dataset in order to train and test new machine learning techniques [4].

In this exercise, we will design and construct a fuzzy expert system that classifies a webpage into four fuzzy categories: safe, weakly suspicious, strongly suspicions and phishing. In addition to the category class, we want to get a numerical score in the range of 0..100 to indicate the level of "phishing risk".

You should take the papers give below as starting point for this exercise. Of course, we will do a very simplified detection system, as we're interested in learning the procedure rather than building a real world tool.

Resources

[1] Dou et al, A Systematic Review of Software-Based Web Phishing Detection, IEEE Communications Surveys & Tutorials, 19(4), 2797-2819 (2017)

[2] R. Zieni, L. Massari and M. C. Calzarossa, Phishing or Not Phishing? A Survey on the Detection of Phishing Websites," in IEEE Access, vol. 11, pp. 18499-18519, 2023, doi: 10.1109/ACCESS.2023.3247135.

[3] Abdelhakim Hannousse, Salima Yahiouche, Towards benchmark datasets for machine learning based website phishing detection: An experimental study, Engineering Applications of Artificial Intelligence, Vol 104, 2021, 104347, https://doi.org/10.1016/j.engappai.2021.104347

[4] UCI dataset: https://archive.ics.uci.edu/ml/datasets/Phishing+Websites# (look at Data Folder -> Phishing website features)

The tasks to be done are:

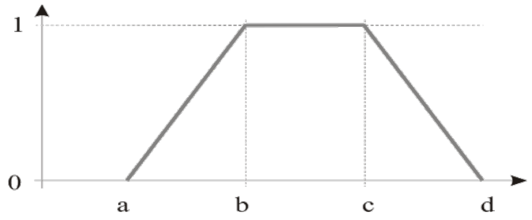
TASK 1: Definition of the input and output linguistic variables.

In paper [3] authors identify 87 features. Select a subset of 5 them and represent them as input linguistic variables (which must take different reference scales and units).

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

As indicated above, for the output variable, you must define four fuzzy categories: safe, weakly suspicious, strongly suspicions and phishing. In addition to the category class, we want to get a numerical score in the range of 0..100 to indicate the level of "phishing risk".

You must use the triangular and 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 definition and analysis.

TASK 3: Implement the fuzzy expert system using Matlab Fuzzy Toolkit.
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: Test the system with four different websites (extracted from the datasets or invented). You must define 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 testing case with screenshots and explanations that justify the output obtained (i.e. showing the activations of rules).

TASK 5: Design (just graphically, no implementation) a more complete fuzzy expert system that includes more features about the websites. Show in a figure the inputs, outputs, and rule blocks that you propose for such expert system. No specific definition of variables nor rules is required.

Submission and deadline:

- A detailed report of the work done in the 5 previous tasks. Report must include all the details about the expert system done in Tasks 1-3 (definitions of variables, rules, screeshots, etc.), about the testing in Task 4 and the answer of Task 5. Matlab files will not be checked during evaluation. Max number of pages should not exceed 20.
- Deliver the Matlab files of your systems too, as validation of the work done and for revision if needed.
- Deadline: **15/12/2024**
- **Only one student of the team uploads the files in the URV virtual campus.**

Evaluation guide:

- TASK 1: definition of variables (correctness, motivation and justification) => 25 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 => 25 points
- TASK 5: advanced design for FES => 10 points