Computational Intelligence (CI-MAI) FL exercise

The energy consumption used for domestic purposes in Europe is, to a considerable extent, due to heating and cooling. This energy is produced mostly by burning fossil fuels, which has a high negative environmental impact. The characteristics of a building are an important factor to determine the necessities of heating and cooling loads. Therefore, the study of the relevant characteristics of the buildings, regarding the heating and cooling needed to maintain comfortable indoor air conditions, could be very useful in order to design and construct energy-efficient buildings.

In this exercise you are going to study the performance of ANFIS models when used to predict the heating load. The data set that you are going to use can be download from the UCI repository in the following link: https://archive.ics.uci.edu/datasets, under the name *Energy efficiency* data set. In the attached article "Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools" (A. Tsanas and A. Xifara), you can find the description of the variables that make up the data and their characteristics.

The idea is that you study the performance of different parameter configurations for an ANFIS model. To compute and report the performance of each configuration and to be able to compare your best results with those of the Tsanas and Xifara paper, you are required to use, at least, the mean absolute error (MAE), as described in the paper. Try to follow as much as possible the methodology described in section 3.3 of the paper to make a fair comparison. However, it should be noted that the run time required to train the ANFIS model can be costly (about 7h to perform a unique run for a whole 10-fold CV), therefore, reduce the CV folds as much as you consider necessary, as well as the number of runs of the training/testing process.

If we were to perform a feature selection study, we would find that the variables Relative Compactness and Glazed Area are the most relevant for predicting the heating load. Knowing this, in this exercise we also want to investigate if it is possible to obtain an ANFIS model using only these two input variables that gives comparable results to the best ANFIS model found in the previous analysis using the eight input features.

Important information:

1) Write a brief document (five sheets maximum) that includes:

- Description of the model configurations and runs that you have performed.
- Explain how you have selected the different parameters and the criteria by which you have selected them.
- Those tables that you consider necessary to describe the results obtained. Explain and reason the results presented in the tables.
- Compare the results you get with those presented in the Tsanas and Xifara article.
- Your own discussions and conclusions with respect the results obtained.
- Include the .m files with your code.

2) To be done in groups of 2 students.

3) Delivery date: no later than January 15, 2024, via the Racó (Practicals).

4) Please include all names in the final document and upload only one document per group.

5) **If you use ChatGPT (or another similar tool) in the document, indicate it every time it is used. We want to evaluate your work, not someone else's.**

NOTE: The exercise is designed to be done in Matlab, but you can implement it in another environment if you like.