GENETIC FUZZY SYSTEMS LAB CLASS GUIDE

Software preparation

Objective: Obtain and execute the KEEL software

1) Web page of the KEEL software: http://sci2s.ugr.es/keel/index.php
2) We will use the Keel software-2014-01-29.zip (There are new versions of the software but are slower)
3) The KEEL software tool includes a large number of different algorithms. You can find a classification summary of all the algorithms included in: http://sci2s.ugr.es/keel/algorithms.php
4) You have available a User Manual: KeelManual2.0.pdf
5) You do not have to download the software from the Keel webpage because you already have it in the zip file that I sent you (*dist* folder). In the folder *dist* you can find the GraphInterKeel.jar, that you can execute directly with a click.

Task 1 – Getting started

Objective: In this task you will learn how to use the KEEL software tool. To this end you will use the *educational* mode of the tool.

1) If you run the KEEL software, you will get the following screen:

As the user manual explains, we can distinguish three parts in the graphic environment:

- The Preparation of the Data Bases part allows users to create different partitions of his own data bases or the data bases available in the KEEL web. Also, it is possible to edit, apply transformations, generate DataSets in the correct format from C4.5 files or view detailed plots about a concrete DataSet.

- The Design of Experiments part has the objective of designing the desired experiments using a graphical interface. After the experiment is designed, the interface generates a .ZIP file containing a directory structure with all the necessary files needed to run those experiments in the local computer.

- The interface also allows the user to add its own algorithms to the experimentation being designed. The only requirement is to accept the input file format of KEEL. Even, it is not needed to use the Java language for the own algorithms of the user. This provides a very flexible way for the user to compare its own methods with the ones in KEEL.

2) In this Lab we will work with the *Educational* tool, in order to see graphically how to define an experiment and the results obtained. From pages 65 to 82 of the User Manual you can find a precise explanation, step by step, about how to perform an experiment in the *Educational* mode and all the menus available.

However, remember that you can perform experiments by means of the graphical interface (*Experiments* option in the main KEEL screen) and run them in a local computer without graphical interface. This is what I recommend you to do when you perform the practical work, because you have much more algorithms available. How to run the experiment that you have designed in the graphical environment is explained in detail in the User Manual.

Task 2 – Classification experiment

Objective: Use a classification example in order to get familiar with the KEEL software tool and analyze the classification results when SLAVE is used. We will chose the Iris benchmark. The data set consists of 50 samples of each of the three species of *Iris* (Iris setosa, Iris virginica and Iris versicolor). Four features were measured from each sample: the length and the width of the sepals and petals, in centimeters.

1) Once you are in the *Educational* experiment design screen, it is necessary to select the type of experiment and the type of partitions to employ; the options selected will determine the kind of methods and data sets that will be available to design the experiment.

Note: 5x2 cross-validation, means 5 times 2 fold cross-validation.

$$\boxed{\begin{array}{ll} \mathcal{T}_1 = \mathcal{X}_1^{(1)} & \mathcal{V}_1 = \mathcal{X}_1^{(2)} \\ \mathcal{T}_2 = \mathcal{X}_1^{(2)} & \mathcal{V}_2 = \mathcal{X}_1^{(1)} \end{array}}$$

$$\boxed{\begin{array}{ll} \mathcal{T}_3 = \mathcal{X}_2^{(1)} & \mathcal{V}_3 = \mathcal{X}_2^{(2)} \\ \mathcal{T}_4 = \mathcal{X}_2^{(2)} & \mathcal{V}_4 = \mathcal{X}_2^{(1)} \end{array}}$$

$$\vdots$$

$$\boxed{\begin{array}{ll} \mathcal{T}_9 = \mathcal{X}_5^{(1)} & \mathcal{V}_9 = \mathcal{X}_5^{(2)} \\ \mathcal{T}_{10} = \mathcal{X}_5^{(2)} & \mathcal{V}_{10} = \mathcal{X}_5^{(1)} \end{array}}$$

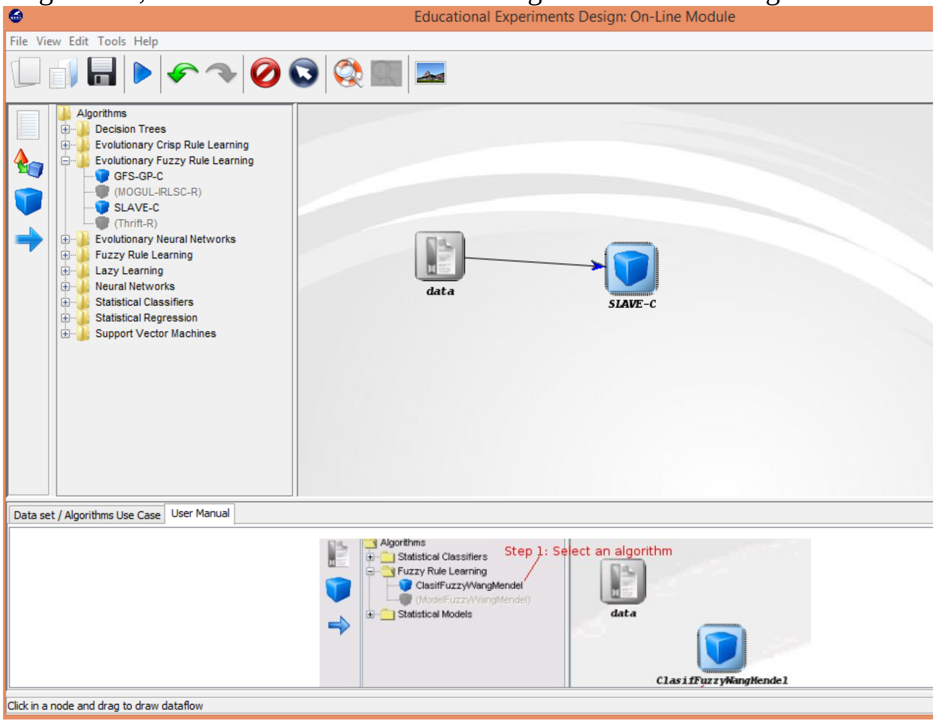
In this experiment select 5-fold cross validation and classification experiment.

2) Select the Iris data set and click on the drawing panel (right hand side of the screen). Notice that in the *status* bar (bottom of the window), the system offers information about the actions that you need to take in each moment in order to perform the experiment.

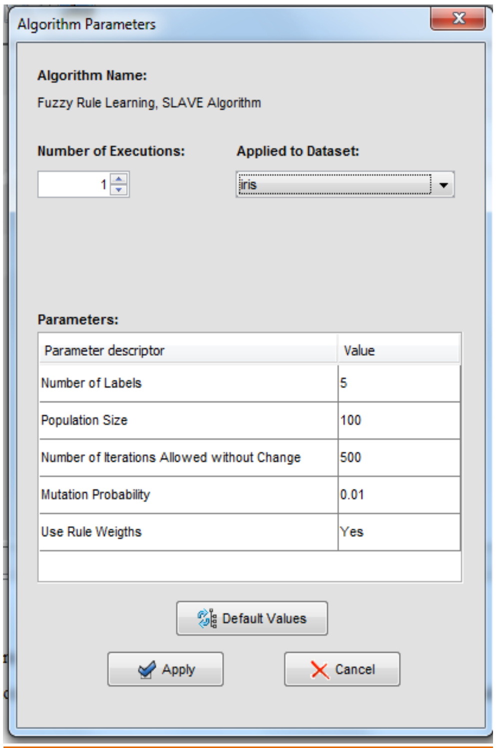
3) Now, it is necessary to select the algorithm that we want to work with. Select the blue box in the left hand side to see the different families of algorithms available in the *Educational* mode. Chose Evolutionary Fuzzy rule Learning, click on the SLAVE-C and click on the drawing panel.

Note: If required, it is possible to include in the experiment design the use of preprocessing algorithms, as for example feature selection, instance selection or discretizers algorithms. When the *Experiments* mode is used instead of the *Educational* one, it is recommendable to include in the experiment design a postprocessing in order to show clearly the results and get some statistics of the performance of the model. In the *Educational* mode the results are shown directly in the GUI.

4) The next step is to set the connections between the different modules. To do this you should select the blue arrow button from the left hand side of the screen. Then, click on the source node and finally click on the target node, i.e. from the data block to the algorithm block. You get the following screen:

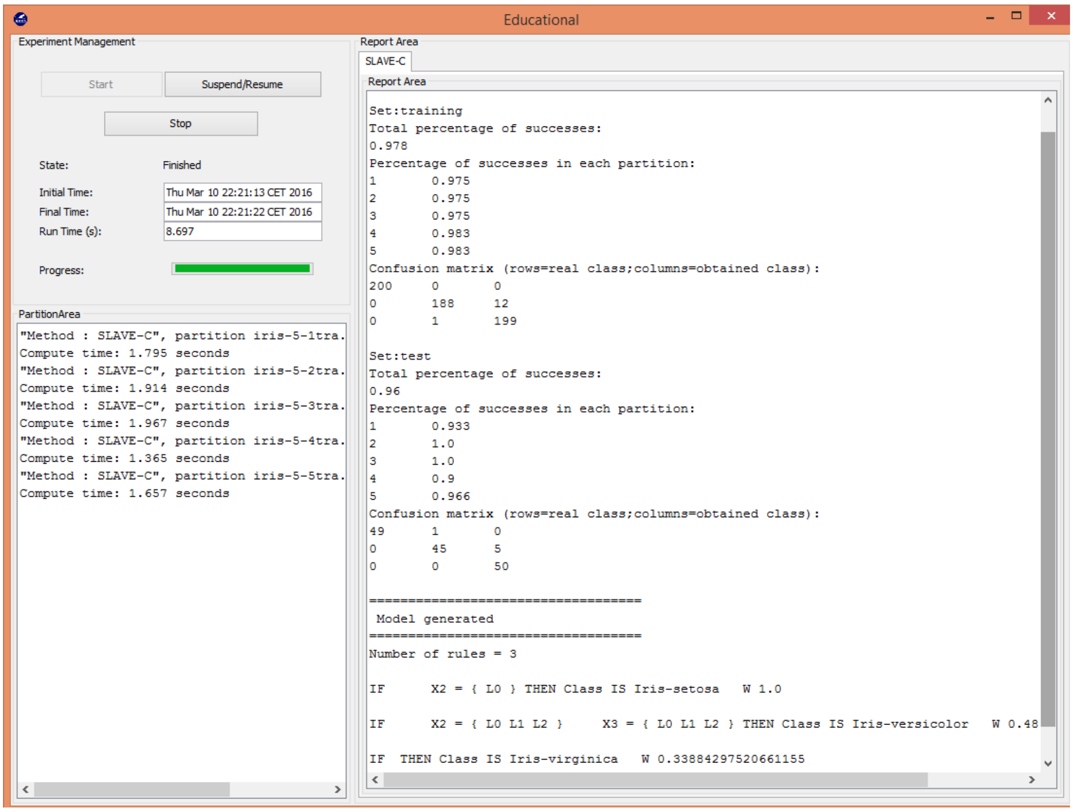


It is possible to take a look to the algorithm parameters by clicking with the right button of the mouse on the algorithms block. Then, you will get the following screen:



In this screen it is possible to change the parameters and to set the number of times that the algorithm will be executed (only available for random methods).

5) Now the experiment is already designed and it is possible to run it. To do so, click the blue triangle at the menu bar. Then, a run screen will be open. You need to push start in order to run the experiment. The results will appear in the screen report area. You should get results similar to the following ones:



The results are also saved in a folder called experiment, where you will find the datasets of your experiment, the scripts and executables of the algorithms chosen in your experiment and the results. The results folder contains a report file with the resume of the results obtained and specific files for each partition. For example, the file result0s0e0 and result0s0e1contain the rules and discretization values that the algorithm obtained for the first partition.

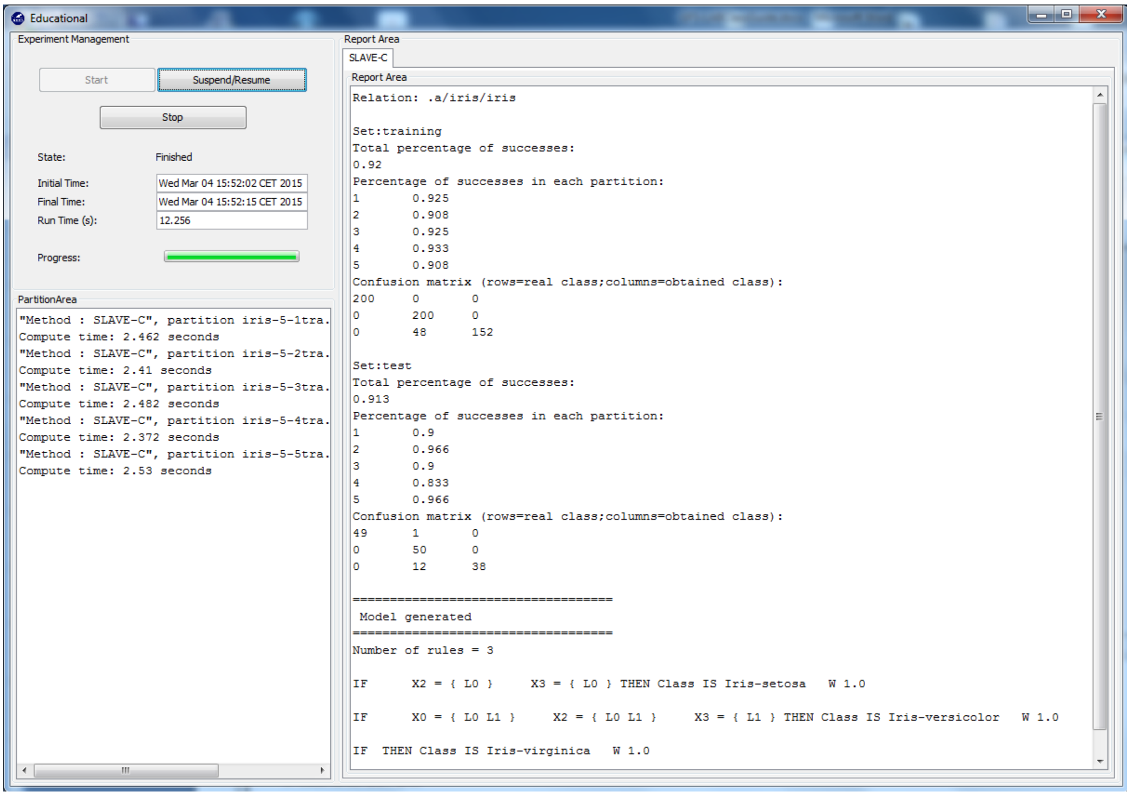
It is possible to save the experiment if wanted.

Task 3 – Play with the parameters

Objective: Change SLAVE parameters to see the changes in the behaviour of the system's model and in the results.

1) Change the different values of the parameters separately and see the robustness of the SLAVE models. Is there any parameter that influence clearly on the accuracy of the SLAVE models for the application at hand?

2) Change several parameters at the same time and see which change implies a reduction of the model's accuracy. For instance, change the number of labels from 5 to 3 and don't use rule weights. In this case the accuracy has been reduced considerably, as can be seen in the following screen.



3) What happens if now the number of labels is set to 9 and rule weights are not used?

4) What happens if the number of labels is set to 9 and the population size is increased to 1000?