

University of Belgrade
Faculty of Organizational Sciences
Center for Business Decision Making

WHIBO GDT Example Process

Belgrade, November 2013

WHIBO generic decision tree (GDT) operator GUI manual

WhiBo generic decision tree user interface contains four panels:

Left panel contains an array of buttons. Every button represents a concrete sub-problem for a decision-tree algorithm design.

Central panel contains:

- Available RCs of selected sub-problem from the left panel.
- Available parameters (if available) for selected RCs.
- Buttons for including or disabling a RC from the current decision tree structure.

Right panel shows current state of user designed algorithm (saved sub-problems, RCs and parameters).

Top panel contains options for creating new, saving current or opening existing generic decision tree algorithm.

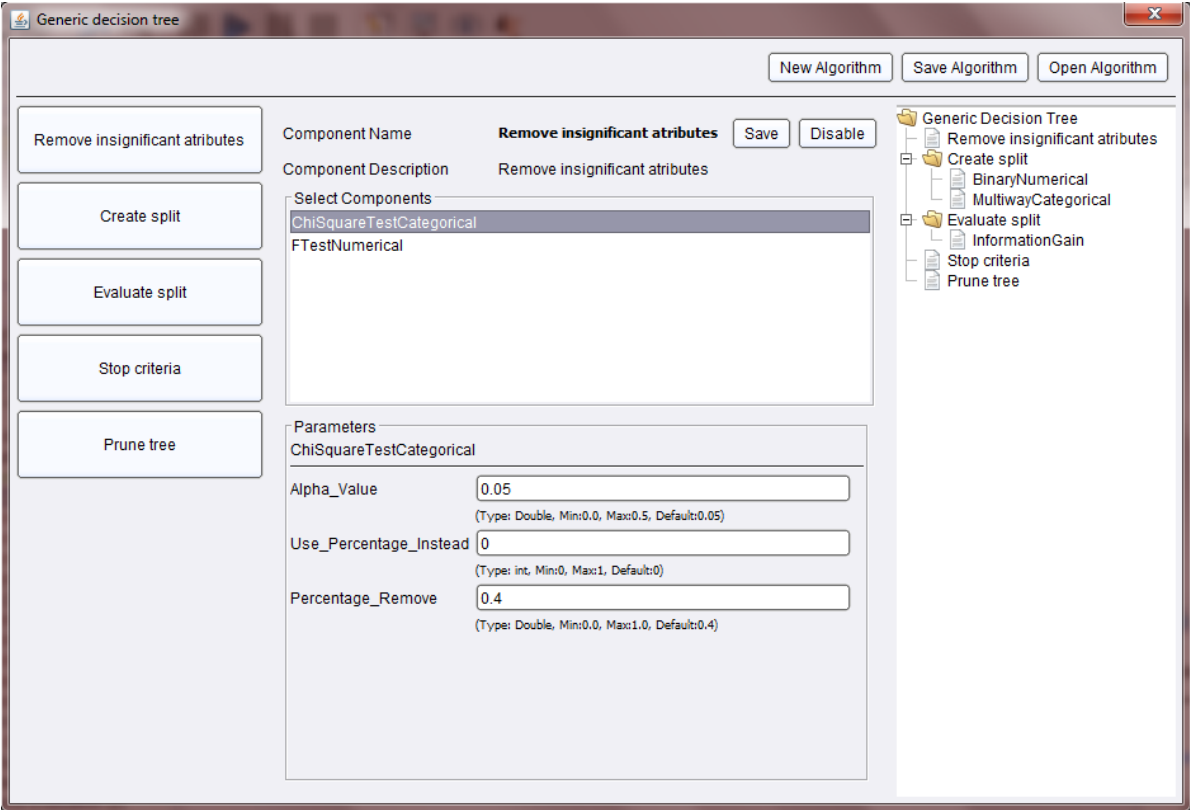


Figure 1 - WhiBo GDT user interface for design of decision tree algorithms

General procedure for designing new algorithms:

- Select sub-problem from left panel. When sub-problem is selected, possible solutions (RCs) are shown in central panel.
- Select RC (or components if multiple) for sub-problem from central panel. If parameters for component(s) are available, they will be shown in bottom part of central panel with their default values.
- Click on save component button. Components and defined parameters for selected sub-problem will be shown in the right panel as part of current GDT algorithm.
- This procedure should be repeated for every sub-problem (Create split and Evaluate split sub-problem are basic for decision tree growth and they must be defined. Definition of other sub-problems is optional). When all sub-problems, components and parameters are defined algorithm should be saved on file system (click on save button from upper panel). By default algorithms are saved with *.wba* (white box algorithm) extension.

WHIBO example process

1. Import dataset into main RapidMiner process. For this purpose Iris dataset will be imported. It can be found in Samples RapidMiner repository in data folder.

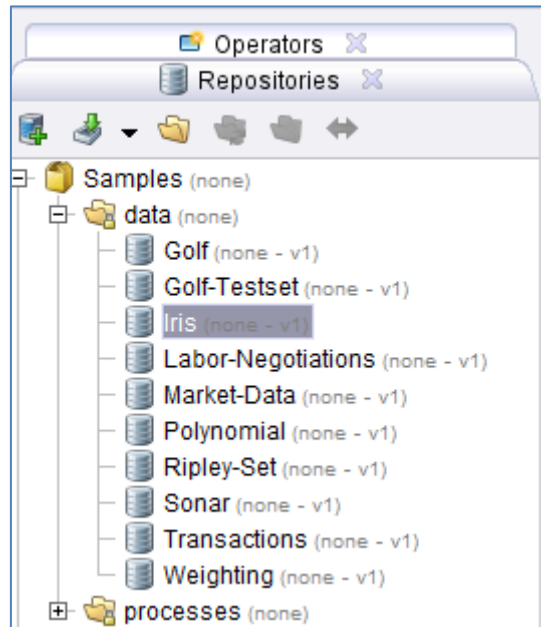


Figure 2 – Sample Repository

Drag Iris dataset into main RapidMiner process.

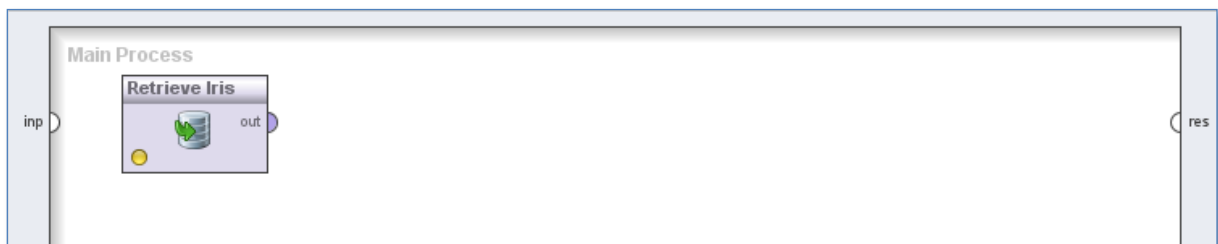


Figure 3 – Main RapidMiner process (1)

2. Import **Generic decision tree** operator.

Generic decision tree operator can be found inside *Operators* panel, in *Whibo* folder.

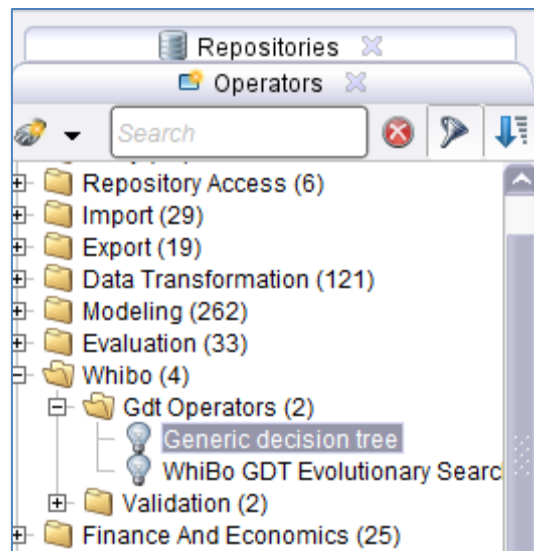


Figure 4 – Generic decision tree operator

Click on this operator and drag it into main RapidMiner process.

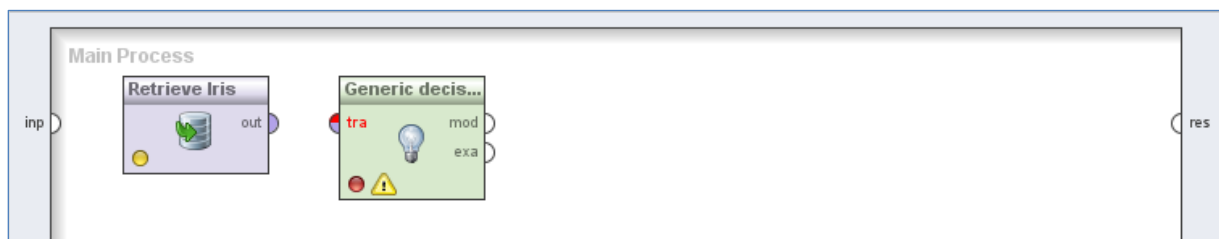


Figure 5 – Main RapidMiner process (2)

When operator is inserted it needs to be connected with dataset. Click on *out* circle on the right side of *Retrieve Iris* operator and connect it with *tra* circle on the left side of *Generic decision tree* operator.

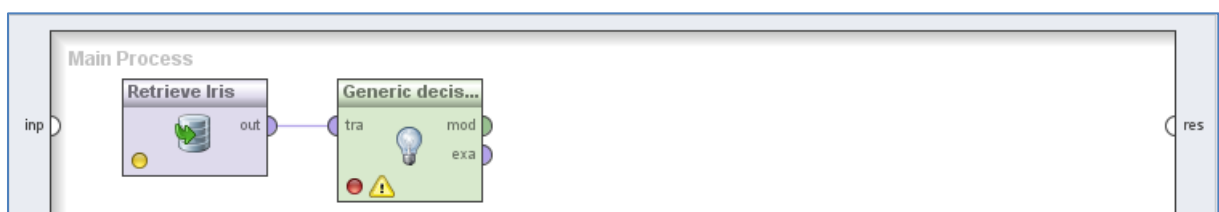


Figure 6 – Main RapidMiner process (3)

3. Click **Design Algorithm** button.

On the *Parameters* panel of *Generic decision tree* operator there is **Design Algorithm** button.

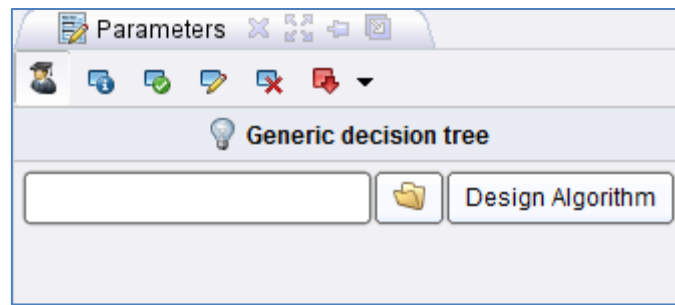


Figure 7 – Parameters panel of *Generic decision tree* operator

When **Design Algorithm** button is clicked new dialog is shown. This is dialog where new algorithm is created. By following general procedure (described above) more than 150.000 of algorithms can be created.

4. Create decision tree by following general procedure.

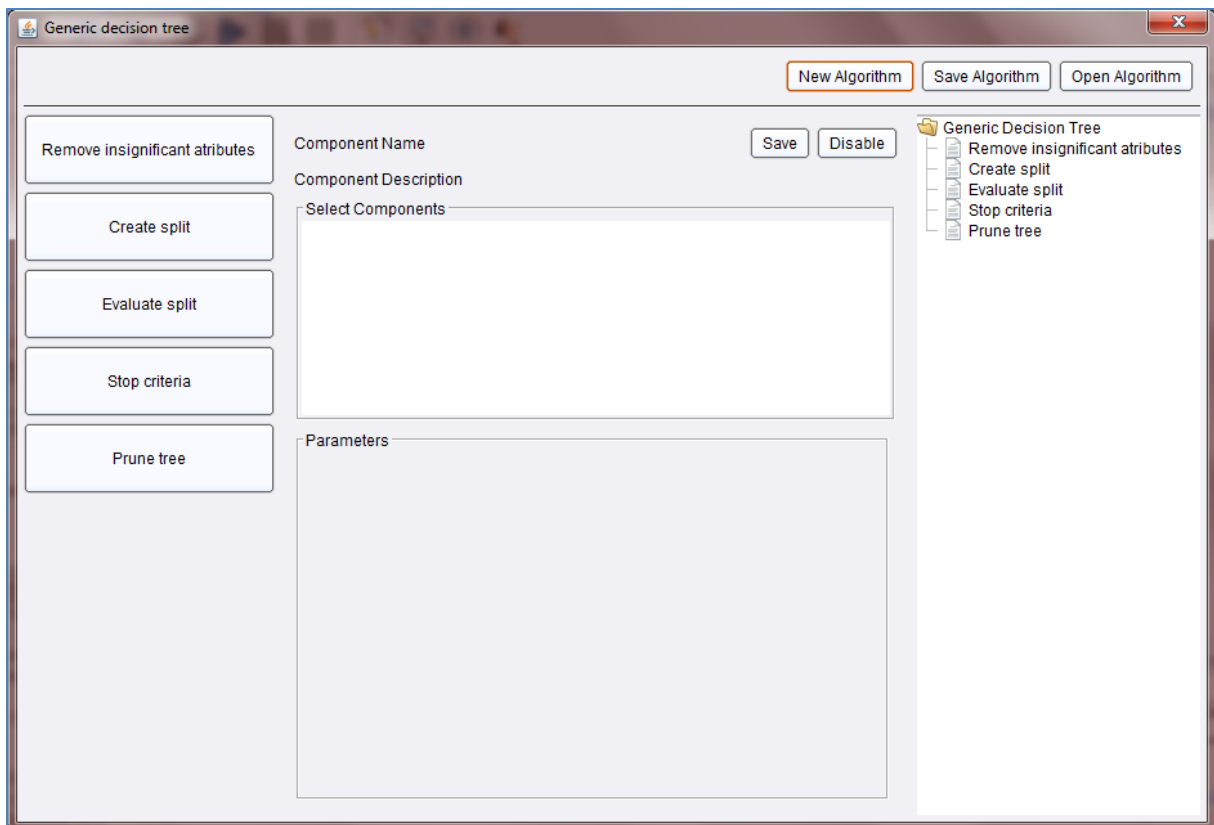


Figure 8 – Algorithm definition dialog

For example CART algorithm will be created.

First, define Create split sub-problem:

- Click on Create split sub-problem on the left panel.
- Select *BinaryNumerical* and *BinaryCategorical* components from central panel (multiple components for one sub-problem are selected by holding CTRL key and clicking on components).
- Click on *Save* component button from central panel.

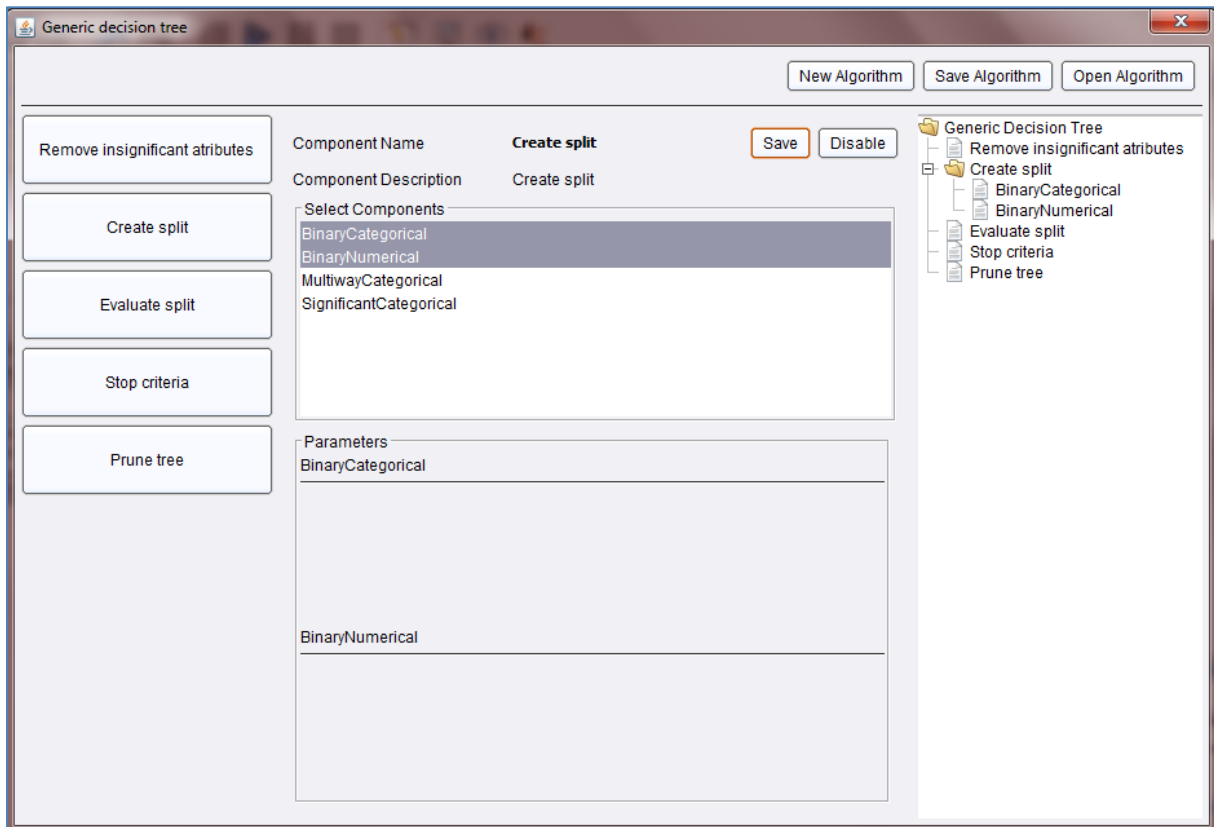


Figure 9 – Create split component of CART algorithm

On the right panel defined components for a sub-problem is visualized through a Tree view.

Next step is definition of evaluate split sub-problem.

- Click on *Evaluate split* sub-problem.
- Select *Gini index* component.
- Click on *Save* component button.

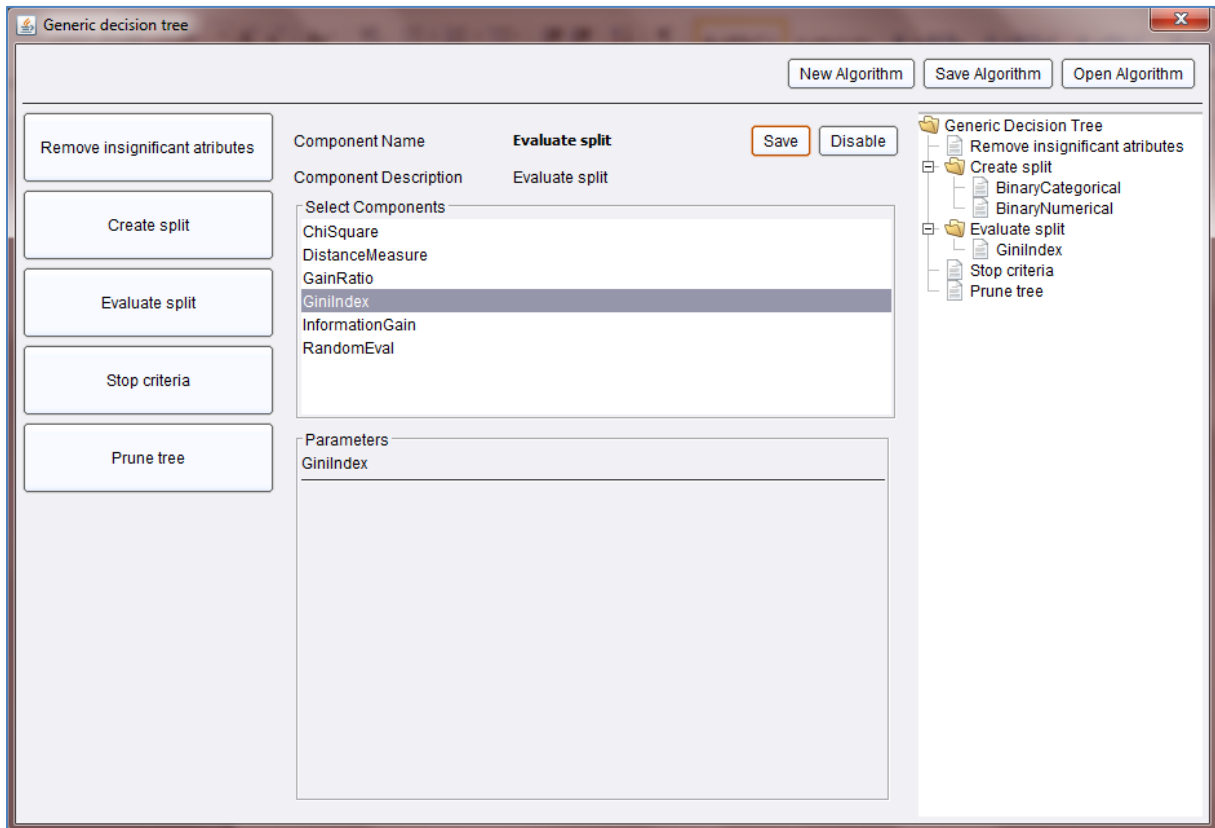


Figure 10 – Evaluate split component of CART algorithm

After definition of algorithm click on **Save Algorithm** button and save decision tree on local hard drive. That (or any other decision tree) should be imported into RapidMiner by clicking **Folder** button in *Parameters* panel and selecting it.

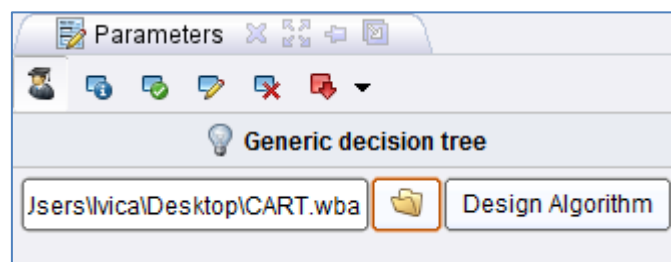


Figure 11 – Folder button

When algorithm definition file is selected its file path will show in text field on left side of *Parameters* panel.

5. Check process for any errors.

In this case only thing that has to be done is to make sure that process is well connected. It should look like one on Figure 12.

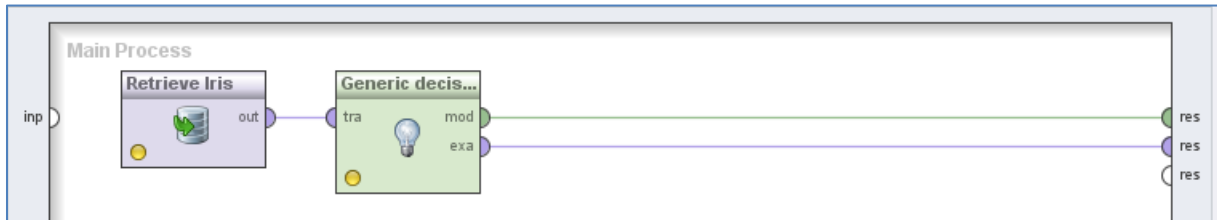


Figure 12 – Main RapidMiner process (4)

6. Click **Run** button.

Process will execute and result will show.

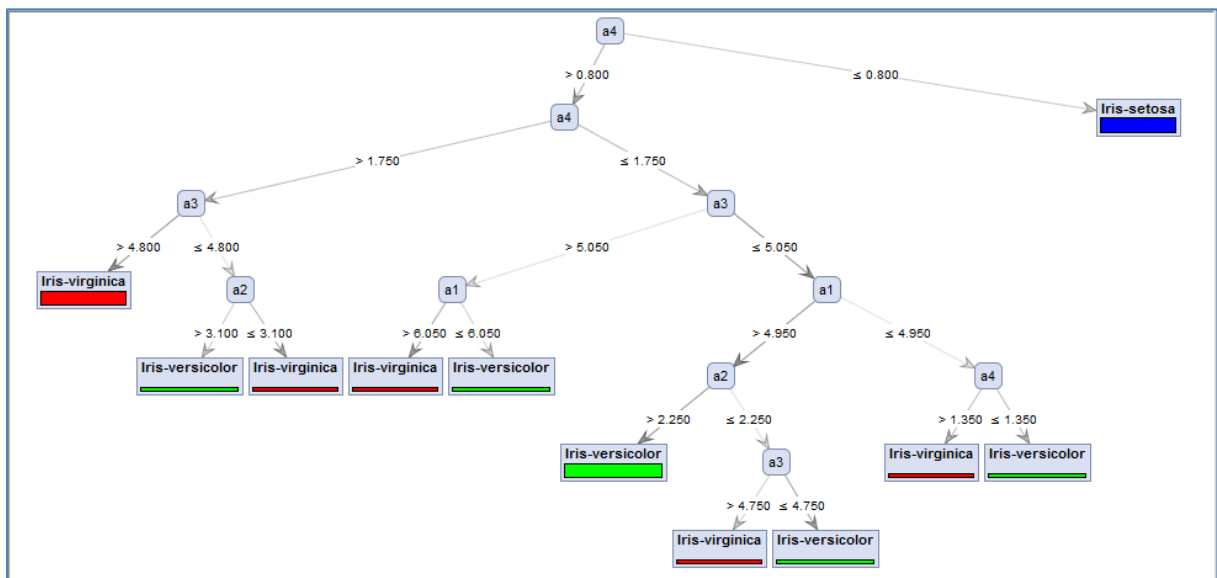


Figure 13 – Decision tree model

For any further information feel free to send e-mail ([contacts](#)) or post on [forum](#).