Prepared By: Divneet Singh Kapoor & Kiran Jot Singh

Worksheet

EXPERIMENT – 9

Name- YANA SRIVASTAVA

UID-20BCS2279 SECTION – 23 B

Aim:

Predict whether or not it will rain tomorrow by training a binary classification model.

Requirements:

(i) Rattle
(ii) R Data Miner.

Expected Outcome:

Rattle specifically uses a simple tab-based concept for the user interface, capturing a work flow through the data mining process with a tab for each stage.

• A typical work flow progresses from the left most tab (the Data tab) to the right most tab (the Log tab).

• For any tab the idea is for the user to configure the available options and then to click the Execute button (or F2) to perform the appropriate task.

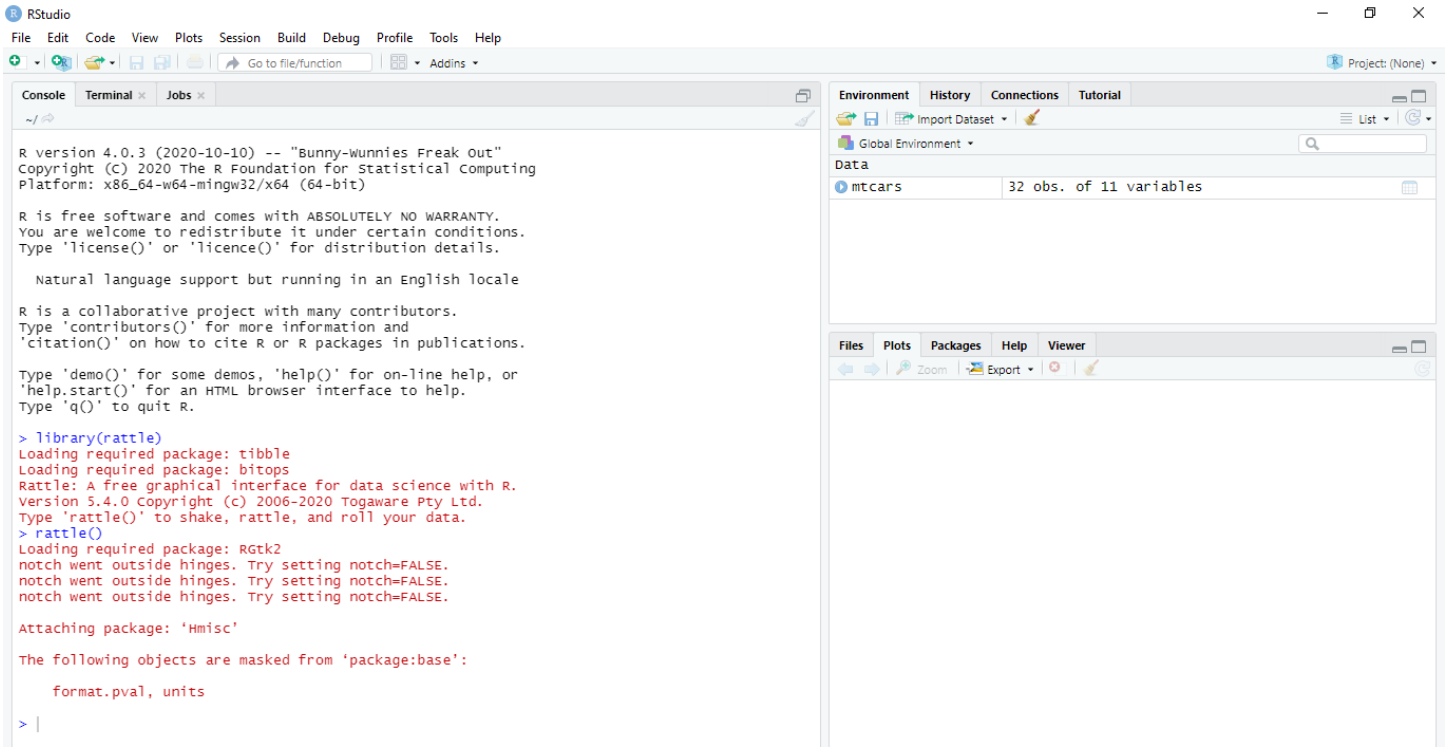
• The status bar at the base of the window will indicate when the action is complete.

• Once we have processed our data, we are ready to build a model and with Rattle we can build the model with just a few mouse clicks. Using a

sample dataset that someone else has already prepared for us, in Rattle we simply:

Steps

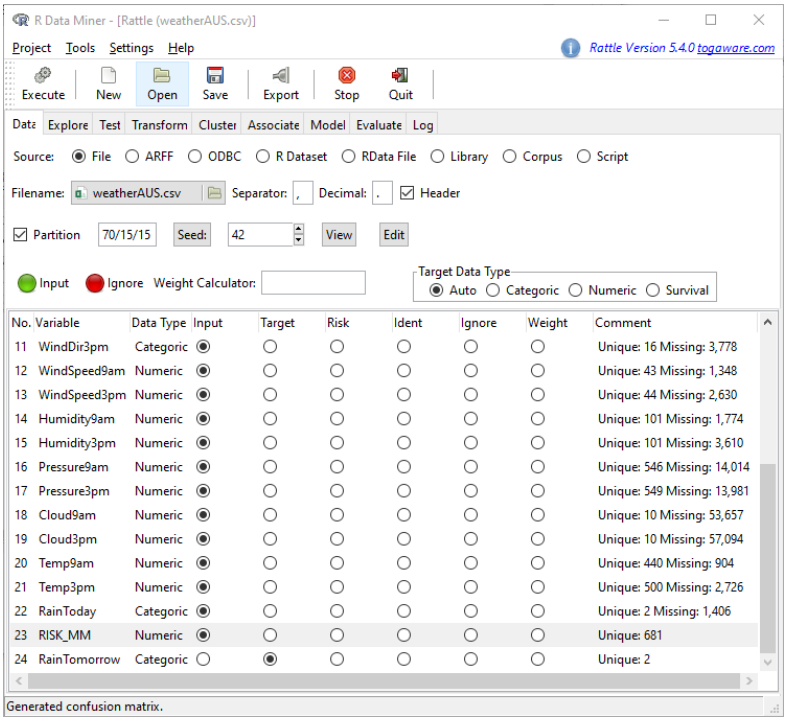
Open the RStudio window and execute the command library (rattle), rattle() to get the R Data Miner window.



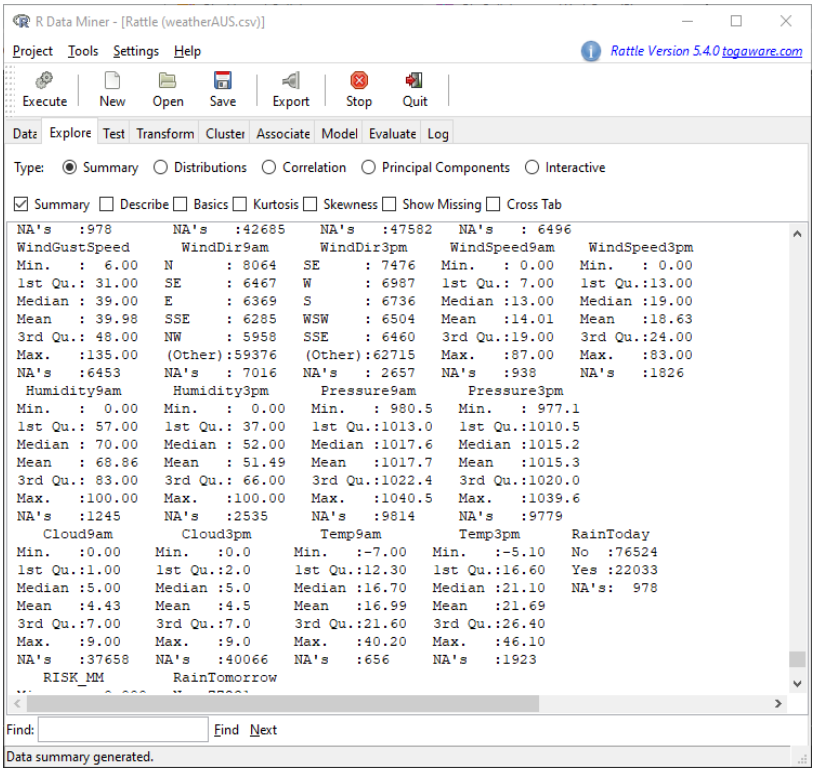
- After the R Data Miner window opens, download weatherAUS.csv data file from any browser and under the filename tab, locate and open the file. The data will be visible

- Now select Rain Tomorrow as target. After selecting the target, execute the data. Every time we make changes like changing the

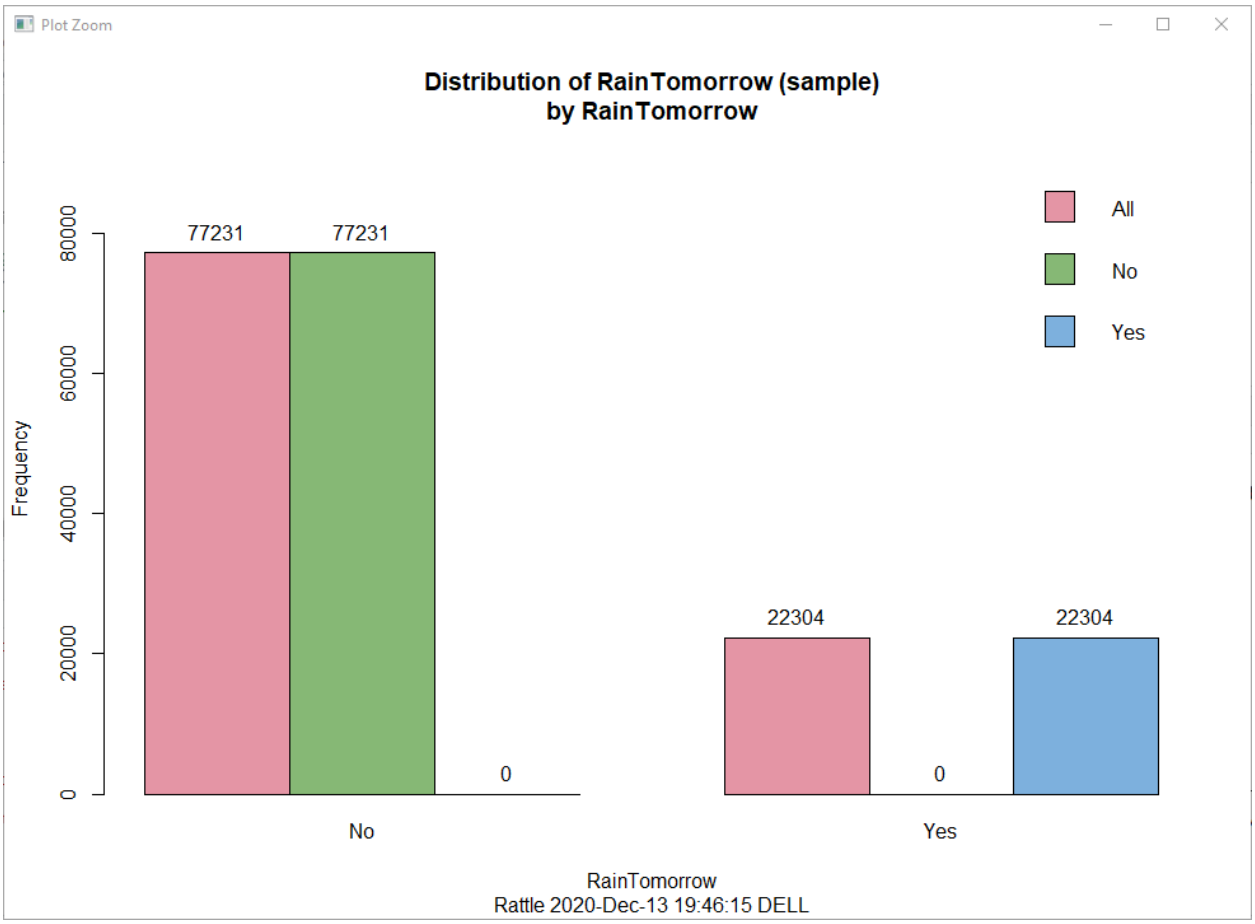
targets, inputs, risks etc. execute the data so that the changes are implied.



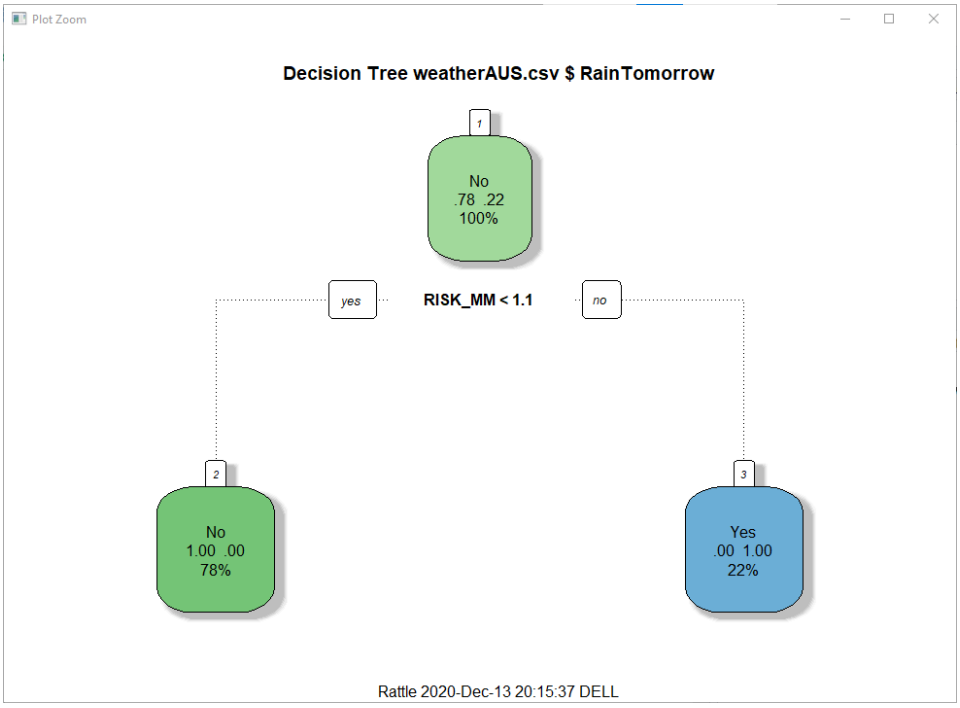
- For the summary statistics, under the data tab select the required variables as inputs. Execute the data. Under the explore tab, select summary from type and execute the selection. The summary of the given data is :



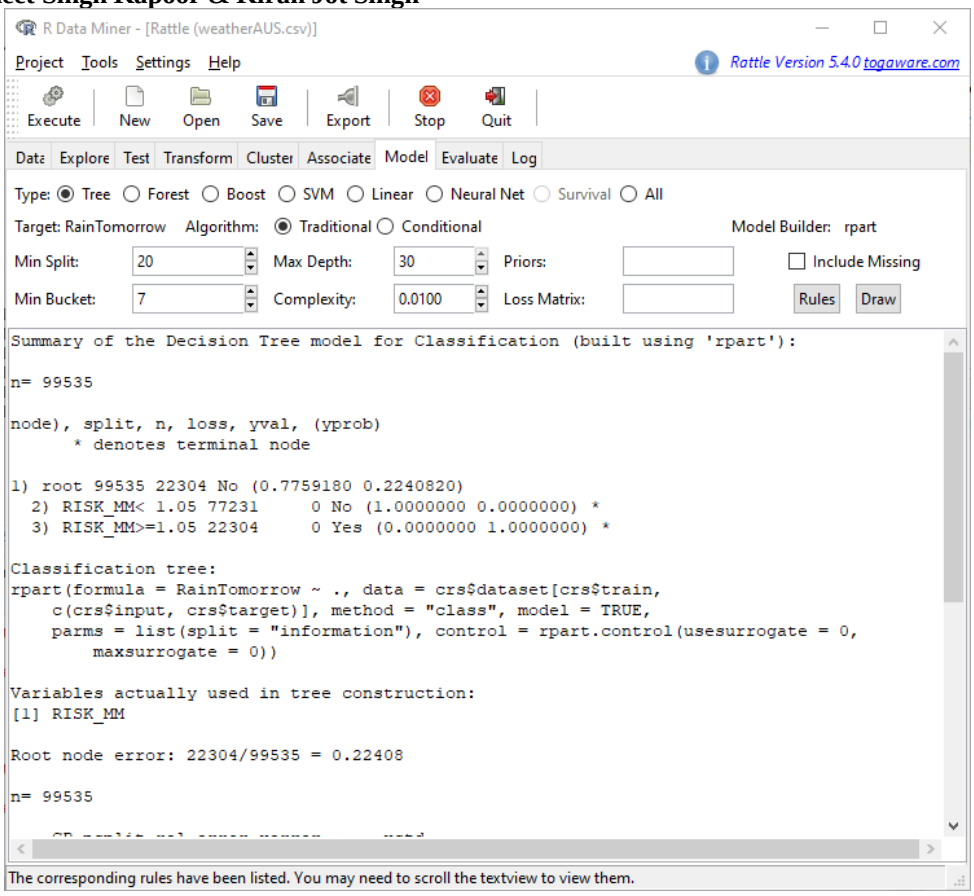
- To get the bar plot for Rain Tomorrow, select Rain Tomorrow as a target. Under the explore tab, select distribution from type and tick on bar plot for Rain tomorrow, execute the selection. The graph will be visible on the RStudio plot area. Click on zoom to view the enlarged plot.



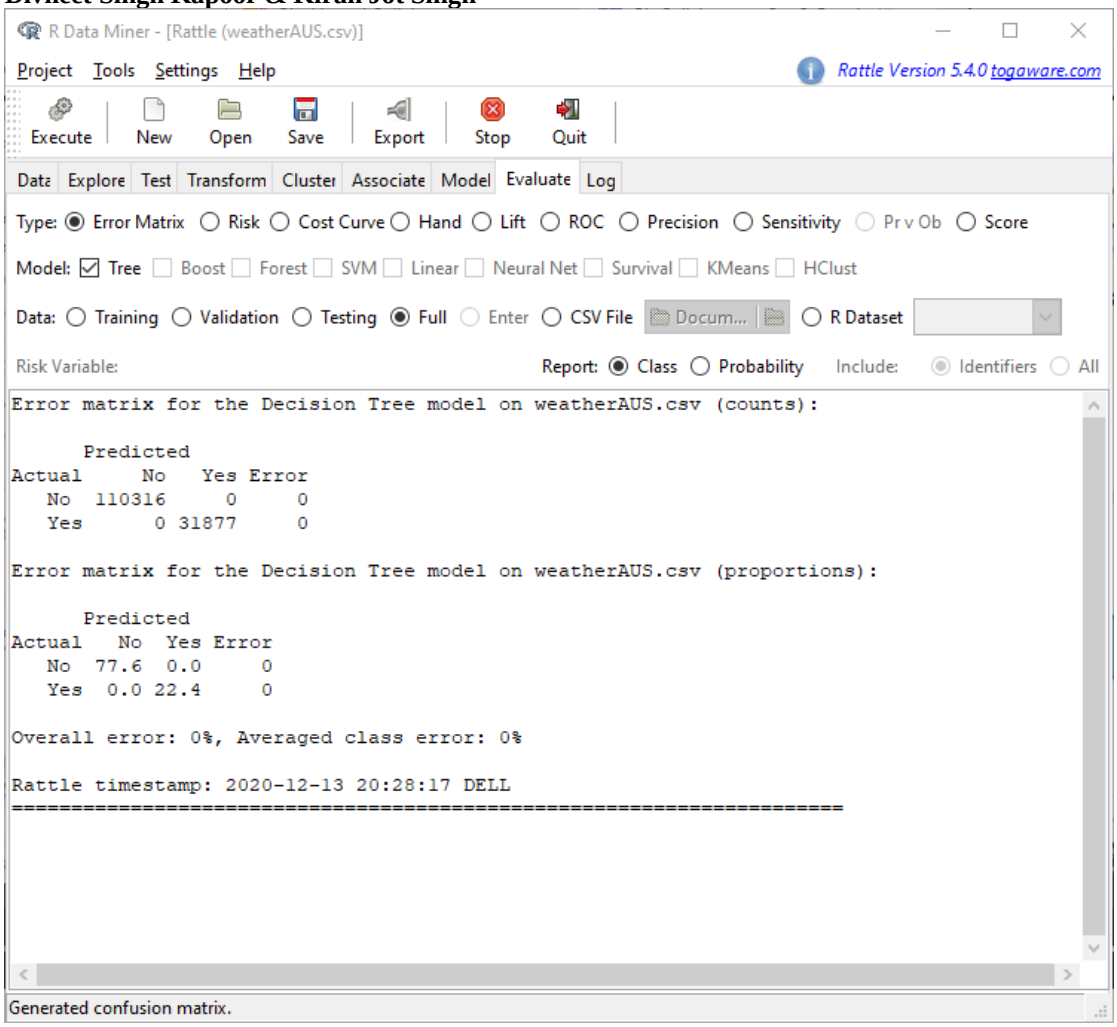
- To get the decision tree for Rain Tomorrow, select Rain Tomorrow as a target. Under the model tab, select tree from type and execute the selection. Two options i.e. Rule and Draw will be visible in the target window, select draw and again execute the selection. The decision tree will be visible in the RStudio plot area. Click on zoom to view the enlarged plot.



- To get the rules for which the decision tree of Rain Tomorrow is made, select Rain Tomorrow as a target. Under the model tab, select tree from type and execute the selection. Two options i.e. Rule and Draw will be visible in the target window, select rule and execute the selection again. The rules will be visible in the R Data miner window.



- To check the correctness and errors in the data, under the evaluate tab, select error matrix from type. Also select full from data tab to check errors in whole data and execute the selection. The errors will be visible in the R Data window.



CONCLUSION:

Rattle continues to undergo development, extending in directions dictated by its actual use in data mining and from suggestions and code offered by its user population.

Thank You Mam

Prepared By: Divneet Singh Kapoor & Kiran Jot Singh