**Experiment 9: Predict whether or not it will rain tomorrow by training a binary classification model.**

**Requirement:** Laptop or Desktop with Python installed

**Theory**

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:

1. Click on the Execute button.

Rattle will notice that no dataset has been identified, so it will take action, as in the next step, to ensure we have some data.

2. Click on Yes within the resulting popup.

The weather dataset is provided with Rattle as a small and simple dataset to explore the concepts of data mining.

3. Click on the Model tab.

This will change the contents of Rattle's main window to display options and information related to the building of models. This is where we tell Rattle what kind of model we want to build and how it should be built.

4. Click on the Execute button.

Once we have specified what we want done, we ask Rattle to do it by clicking the Execute button. For simple model builders for small datasets, Rattle will only take a second or two before we see the results displayed in the text view window.

The data comes from a weather monitoring station located in Canberra, Australia, via the Australian Bureau of Meteorology. Each observation is a summary of the weather conditions on a particular day. It has been processed to include a target variable that indicates whether it rained the day following the particular observation. Using this historic data, we have built a model to predict whether it will rain tomorrow. Weather data is commonly available, and you might be able to build a similar model based on data from your own region.

With only one or two more clicks, further models can be built. A few more clicks and we have an evaluation chart displaying the performance of the model. Then, with just a click or two more, we will have the model applied to a new dataset to generate scores for new observations. Now to the details. We will continue to use Rattle and also the simple command line facility. The command line is not strictly necessary in using Rattle, but as we develop our data mining capability, it will become useful. We will load data into Rattle and explain the model that we have built. We will build a second model and compare their performances. We will then apply the model to a new dataset to provide scores for a collection of new observations (i.e., predictions of the likelihood of it raining tomorrow).

