Download WEKA – Waikato Environment for Knowledge Analysis - Learning Tool

University of Waikato, New Zealand

To download Weka go to https://www.cs.waikato.ac.nz/ml/weka/

Building a NN with Weka

Build a multilayer NN to learn to predict “outcome”

#	Temp	opponent winning %	windy	outcome
1	cool	40	Yes	Win
2	cool	55	No	Lose
3	cool	60	Yes	Lose
4	cool	75	No	Lose
5	cool	30	No	Win
6	warm	85	No	Win
7	warm	10	Yes	Win
8	warm	62	No	Win
9	warm	30	No	Win
10	cold	50	Yes	Lose
11	cold	35	Yes	Lose
12	cold	70	No	Win
13	cold	40	No	Win
14	cold	65	No	Win

Building a NN with Weka

STEP 1:Create an “arff” data file that Weka reads as input for training

```
@RELATION footballtrain
@ATTRIBUTE Temp {cool, warm, cold}
@ATTRIBUTE Opponentpct  NUMERIC
@ATTRIBUTE Windy        {Yes, No}
@ATTRIBUTE Outcome      {Win, Lose}
@DATA
cool,40,Yes,Win
cool,55,No,Lose
cool,60,Yes,Lose
cool,75,No,Lose
cool,30,No,Win
warm,85,No,Win
warm,10,Yes,Win
warm,62,No,Win
warm,30,No,Win
cold,50,Yes,Lose
cold,35,Yes,Lose
cold,70,No,Win
cold,40,No,Win
cold,65,No,Win
```

Building a NN with Weka

Step 2: Loading the file and selecting options

1. Click on explorer on the Weka chooser window.

2. Click on “open file” and select the arff file that was created in step 1.

3. Click on the “classify” option of explorer window

4. Under the “classifier” panel, click on “Choose”

5. Within the window that appears after you click on “Choose” expand the “functions” header

6. Within the window that appears after you click on “functions” click on “Multilayer Perceptrons”

7. Within the “Test options” panel, select “use training set”

Building a NN with Weka

Step 3: Training the NN and seeing results

1. Click on “start” button

2. Training results will appear inside the “classifier output” window.

3. Examine the results inside the “classifier output” window from top to bottom. Identify the following
   - ❑ Accuracy obtained, if not 100% then which records were incorrectly classified.
   - ❑ # of layers, # of neurons in each layer, weights associated with inputs to neurons
   - ❑ Which neurons in output layer are assigned to “win” outcome, and which to “lose” outcome,
   - ❑ Draw the network architecture – with inputs, hidden layers, output layer, weights of connections.

Building a NN with Weka

Step 4: Saving the model

1. Right click on the model that appears under “results list”

2. Select the “save model” option and place it in the directory of your choice.

Changing Building Parameters within WEKA

After opening an arff file, selecting “classify”, choosing “multilayerPerceptron” ….

1. Click on the MultilayerPerceptron instance that appears next to “choose” button.

2. Within the “GenericObjectEditor” that appears, the following choices are useful

   a. GUI – set to true to see the network architecture being used.

   b. hiddenLayers – when set to “a” WEKA will create one hidden layer with
      # of neurons =(#of inputs + # of outputs)2. this can be change as follows – 4,3 will create two hidden layers with 4 neurons in first one and 3 in the second one.

   c. Training time refers to the number of epochs to train the network. In most cases it should be > 10000

Changing Building Parameters within WEKA

3. Train the following networks and check if the accuracy improved from the one built using default parameters

   a. 2 hidden layers with 4 and 3 neurons respectively and number of epochs = 2000 BTW – what is an epoch?

   b. Same as “a” but increase epochs to 10,000

   c. 3 hidden layers with 4, 3, 2 neurons respectively and number of epochs = 2000

An Approach to changing parameters

1. Using Weka “a” default for hidden layer, increase “training time” parameter from 500 to 1000 to 2000 to 5000 to 10000 checking accuracy at each step

2. If “1” is not sufficient then try increasing the number of hidden units by replacing “a” with a specific number. I would first try doubling the number of hidden units . Increase training times as indicated in 1 above.

3. If “2” is not sufficient then try increasing the number of hidden layers from 1 to 2 to 3 and more if needed. It is best if the number of hidden units start larger in the first hidden layer and get progressively smaller as they get to the output layer. The last hidden layer before the output layer should have at least the same number of units as the output layer.

Building a NN with Weka

Step 5: Testing the Model

1. Create an “arff” data file that Weka can read as input for testing – notice the “?” symbol for where the value of “outcome” should be for each record

```
@RELATION footballtest
@ATTRIBUTE Temp {cool, warm, cold}
@ATTRIBUTE Opponentpct   NUMERIC
@ATTRIBUTE Windy         {Yes, No}
@ATTRIBUTE Outcome       {Win, Lose}
@DATA
cool,40,Yes,?
cool,55,No,?
cool,60,Yes,?
cool,75,No,?
cool,30,No,?
warm,85,No,?
warm,10,Yes,?
warm,62,No,?
warm,30,No,?
cold,50,Yes,?
cold,35,Yes,?
cold,70,No,?
cold,40,No,?
cold,65,No,?
```

Building a NN with Weka

Step 5: Testing the Model - continues

2. Click the “Explorer” button to open the Weka Explorer interface.

3. Load the “.arff file that you created to test the Neural Network.

4. Click the “Classify” button.

5. Right click on the empty “Result list” window pane and click “Load model”

6. Select the neural net model previously saved.

7. On the “Classify” tab, select the “Supplied test set” option in the “Test options” pane.

8. Click the “Set” button. Then click the “Open file” button on the options window and select the file to test the Neural Network.

9. Close the window.

Building a NN with Weka

Step 5: Testing the Model - continues

10. Click the "More options…" button in the test options window.

11. Uncheck the following information:
    - ❑ "Output model"
    - ❑ "Output per-class stats"
    - ❑ "Output confusion matrix"
    - ❑ "Store predictions for visualization"

12. For the "Output predictions" option click the "Choose" button and select "Plain Text".

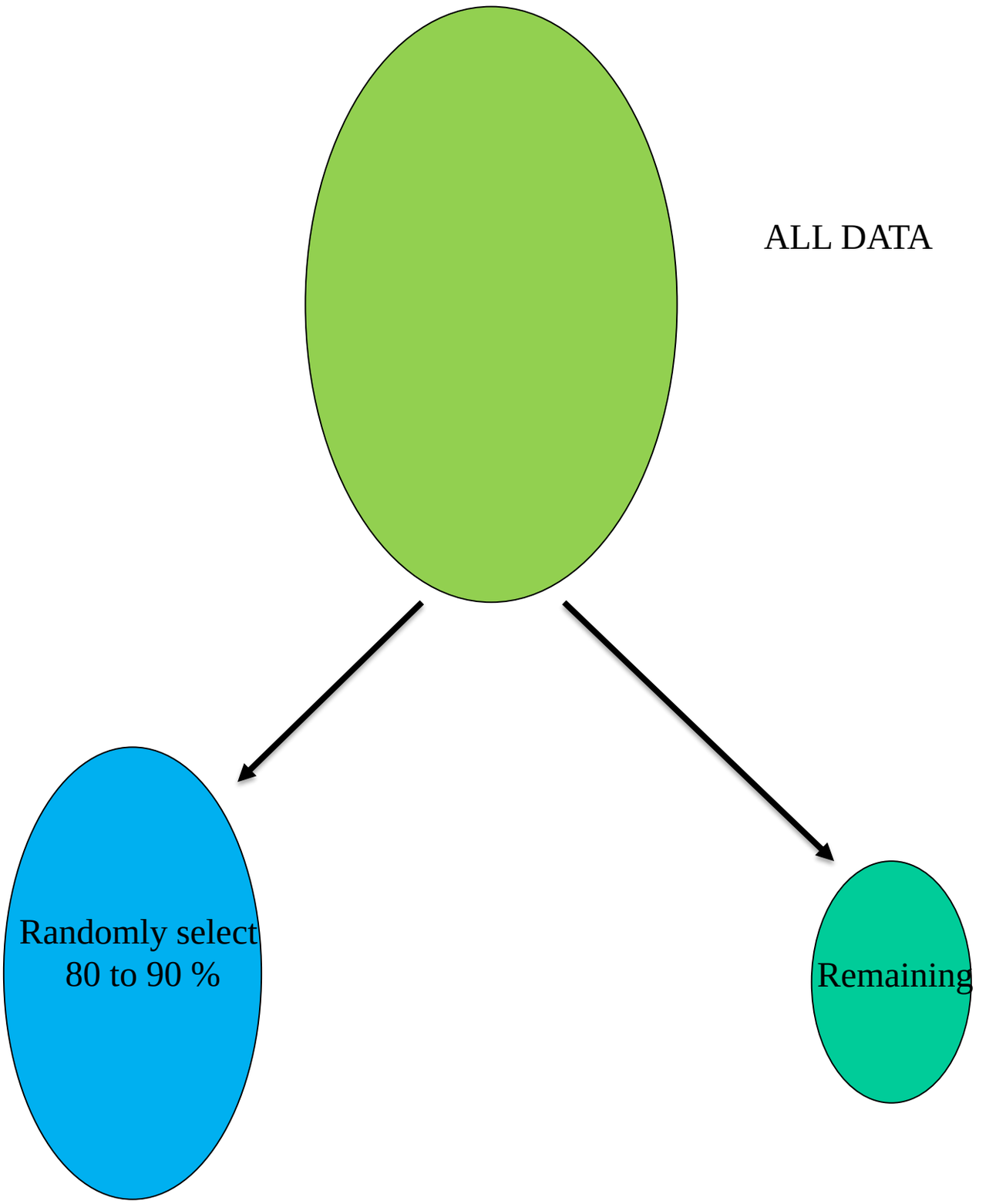
13. Click the "OK" button to confirm the Classifier evaluation options.

14. Right click on the list item for your loaded model in the "Results list" pane.

15. Select "Re-evaluate model on current test set".

16. The predictions for each test instance are then listed in the "Classifier Output" pane.

Creating training and testing files



Creating random samples of a file with WEKA

The following steps show how to sample an “arff” file randomly and create a “training that includes 90% of the records in the original “arff” file and another file “test” that includes the other 10%.

1. “open” the “.arff file that you want to sample with Weka Explorer – see # of instances in “current relation” window

2. Within “filter” window click “choose”.

3. Within window that opens, open “filters”. Within window that opens, open “unsupervised”. Within window that opens, open “instance”. Within window that opens, select “resample”.

4. Click on “resample” instance that now appears next to “choose” inside the “filter” window.

5. Inside the “GenericObjectEditor” window that appears, set “noReplacement” to “true” and “sampleSizePercent to “90” . Click “OK”

6. Click the “Apply” button inside the “filter” window. See # of instances in “current relation” window – reduced by 10% from number in step 1 above.

Creating random samples of a file - continues

7. Save the file naming it “training.arff”.

8. Click on “undo” button. See # of instances go back to original number.

9. Click on “resample” instance that appears next to “choose” inside the “filter” window.

10. Inside the “GenericObjectEditor” window that appears, set “noReplacement” to “true”, “invertSelection” to true. Leave “sampleSizePercent” to “90”. Click the “Ok” button.

11. Click the “Apply” button inside the “filter” window. See # of instances in “current relation” window – reduced by 90% from number in step 1 above.

12. Save the file naming it “test.arff”.

A Broader View of Building Multilayer Perceptrons with WEKA

Part 1 – Assignment 5

Building a NN when:

1. You have selected, beforehand, a test set that is of critical importance

2. Letting the NN building tool control all parameters

    a. Max # of epochs

    b. # of hidden layers

    c. # of neurons in the hidden layers

3. Checking overall training accuracy vs testing accuracy

    a. What do you do if testing accuracy is lower than you want?

4. Checking accuracy of individual neural network outputs

    a. Does it matter?

    b. What if they differ?

A Broader View of Building Multilayer Perceptrons with WEKA

Part 2 – Assignment 5

Building a NN when:

1. You don't have a pre-specified test set and instead let the NN building tool split the data – in this case "data set 1"

2. Letting the NN building tool control all parameters

    a. Max # of epochs

    b. # of hidden layers

    c. # of neurons in the hidden layers

3. The accuracy reported by Weka would be the accuracy on the testing set.

4. Testing the NN as it is being deployed on data that it has never seen before (exemplified by data set 2 in this case)
    a. Overall accuracy

    b. Accuracy of individual outputs

    c. What to do if accuracy is not acceptable?

A Broader View of Building Multilayer Perceptrons with WEKA

Part 3 – Assignment 5

Building a NN when:

1. Similar to Part 2 when you don't have a pre-specified test set and instead let the NN building tool split the data – in this case "data set 1"

2. You will search for the best network based on overall accuracy and accuracy of individual outputs, for data set 1 varying important NN training parameters
   a. Max # of epochs

   b. # of hidden layers

   c. # of neurons in the hidden layers

3. Testing the most promising NNs found in step 2 above on data that it has never seen before (exemplified by data set 2)
   a. Overall accuracy

   b. Accuracy of individual outputs

4. Making a selection for a specific NN and explaining why