RMIT University
School of Science
COSC2110/COSC2111 Data Mining
Laboratory Week 9

Aims of this lab

1. Learn how to develop a bash script to convert a Weka .arff file to a javanns .pat file
2. How to create a new network with JavaNNS
3. How to train a network with training and validation sets and get the accuracy on a test set.

This lab requires the files:
`code-and-scripts/iris-for-javanns.sh`
`data/arff/student-data-lab09.arff`
in the directory
`/KDrive/SEH/SCSIT/Students/Courses/COSC2111/DataMining`

The task is to generate javanns pattern files from the arff file, and train a neural network with these pattern files.

1. Examine the file `student-data-lab09.arff` and determine the necessary data encoding. The output variable is Sex.
2. Load the file into weka, normalize the attributes with a filter and save.
3. Examine the file `iris-for-javanns.sh`. This file takes the iris data and randomly splits it into training, validation and test pattern files.
4. Modify the file `iris-for-javanns.sh` to only generate `student-data-train.pat`, and exit immediately.
5. Using javanns, generate a suitable network and verify that your `student-data-train.pat` can be loaded and be successfully used for training. Instructions for this are on the echo recording of 16-Sep-2019 starting around minute 25.
6. Carry out a training run and get a result file. Use the analyze program to get the classification error rate.
7. Extend the script to generate student-data-valid.pat and student-data-test.pat
8. Repeat the training with all 3 pattern files and get the result file. Instructions for this are on the echo recording of 16-Sep-2019 starting around minute 12.