# 08338 ACW Lab s

This lab sheet follows on from sheets ( **08338-Lab-ACW-Sheet2 )** and ( ***08338-Lab-ACW-Sheet1*** ).

## Overview

This sheet starts after Item 15 from the previous sheet (**08338-Lab-ACW-Sheet2**) and proceeds to cover new items.

## Sheet 3: Decision Rules

Here is the guide to what should achieve from this lab guide to the ACW. Your steps may vary slightly but the principles and process are as follows:

19. In Excel create a new worksheet named **DecisionTables**. Copy the first line of the data files (**Attribute names**) to the first line in this worksheet, including Id. Now map your rules from J48 into the worksheet. Format the rules so that each rule fits with the data description of the first line. Give each rule a unique name (in Id), for example J48-1, etc. The example at the end of this document (Table 1 below) shows how to format j48 rules
20. Now choose a rule classifier (from weka→Classify→Rules) and apply to the same data as used for J48. Your Classifier Performance work should have identified which of these it will be.
21. Save the output to a suitably named worksheet. Map your rules from there into the **DecisionTables** worksheet. Format the rules so that each rule fits with the data description of the original data. Give each rule a unique name (in Id), for example JRip-1 etc.
22. Now look for conflicts and agreements across the two rule sets. Note these in a new worksheet (***ConflictSet***). Examples of potential conflicts are given below in Table 2.
    a. The work on Conflict Sets can be as complete and involved as you wish. The minimum requirement for the ACW is given in the ACW description.
    b. The minimum requirement is to identify first those rules that give **Risk=High** to unclassified data records (the five test data records). Now look for rules that agree (Supportive Conflicts) and those that give Risk=Low (Unsupportive Conflicts) on the test data.
23. Now copy your clean data worksheet to a new worksheet and name it CleanX-Nominal. Convert all numeric attributes (except Id) to nominal. Slides 102 to 110 in the lecture **DMDS-7-ReductionandTransforms** give hints on how to do this. Save as an appropriately named csv file.
24. Load the new nominal data into weka. Use J48 on it to check what has happened to the classifier performance. If the performance is very poor, you may need to modify your numeric to nominal data transforms. Repeat step 23 to 24 until satisfied with the performance of J48 on the nominal data.
25. The Nominal data mapping needs to be documented in **Methodology** and **FinalDataDescription.**
26. Once loaded into weka, the new nominal data can be used with an Association Rule generator (e.g. Tertius). Note you will have to modify the parameters to the Rule generator to get rules that predict Risk; and you may need to increase the number of rules generated.
27. Repeat steps 21 and 22 with the rules from the Association Rule generator. Note the **DecisionTables** worksheet may eventually include a very good number of rules

Questions via lab, lectures, forum or email

| Id | Indication | Diabetes | IHD | Hypertension | Arrhythmia | History | IPSI | Contra | Risk | Info |
|---|---|---|---|---|---|---|---|---|---|---|
| j48-1 | tia | | | | yes | | | <=40 | Low | 6 0 |
| j48-2 | a-f | | | | yes | | | <=40 | High | 14 0 |
| j48-3 | cva | | yes | no | yes | | | <=40 | Low | 2 0 |
| j48-4 | cva | | yes | yes | yes | | | <=40 | High | 12 5 |

```
Arrhythmia = yes
|   Contra <= 40
|   |   Indication = tia: Low (6.0)
|   |   Indication = a-f: High (14.0)
|   |   Indication = cva
|   |   |   IHD = yes
|   |   |   |   Hypertension = no: Low (2.0)
|   |   |   |   Hypertension = yes: High (12.0/5.0)
```

***Table 1: Example showing how to Format j48 Rules into Decision Table in Excel (colour coded)***

| Id | Indication | Diabetes | IHD | Hypertension | Arrhythmia | History | IPSI | Contra | Risk | Coverage | Conflicts |
|---|---|---|---|---|---|---|---|---|---|---|---|
| j48-1 | tia | | | | yes | | | <=40 | Low | 6 0 | Jrip8 |
| j48-2 | a-f | | | | yes | | | <=40 | High | 14 0 | Jrip8 |
| j48-3 | cva | | yes | no | yes | | | <=40 | Low | 2 0 | Jrip3 |
| j48-4 | cva | | yes | yes | yes | | | <=40 | High | 12 5 | Jrip7<br>Jrip8 |
| Jrip7 | cva | | | yes | | | >=90 | | High | 12 5 | J48-4<br>Jrip8 |
| Jrip8 | | | | | | yes | >=80 | | High | 8 0 | J48-1<br>J48-2<br>J48-3<br>J48-4<br>Jrip7 |

***Table 2: Example showing potential conflicts (colour coded) between j48 and Jrip Rules (Decision Table in Excel)***