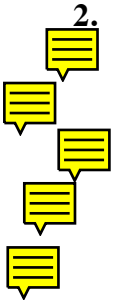
CSCI3230 Fundamentals of Artificial Intelligence
Weka Laboratory Specification
Fall 2013

1. Introduction

In this Weka Laboratory, you are asked to use the data mining tool, Weka, with its embedded algorithms and its limitations to work on a provided dataset. You are required to attend one of the Week 9 tutorials. Tutors will grade students based on their progress and understanding about Weka and data mining in the laboratory sessions. The full mark of this Weka Laboratory is 10, which is 10% of the whole course. Please note that the one who does not attend the tutorials without any prior notice to the tutors will get ZERO marks.

2. Tasks



i) Download the (installation free) Weka program from our course homepage
ii) Download the provided dataset from our course homepage
iii) Convert the dataset from *.names/*.data to *.arff (the format that Weka accepts)
iv) Preprocess your data
v) Perform classification on the dataset using different classifiers.
vi) Evaluate the classification performance
vii) Interpret the overall result

3. Marking Scheme

i) The way you convert the dataset to *.arff format
ii) The way you preprocess the dataset
iii) The way you build classifiers
iv) The way you evaluate your classification performance
v) The way you interpret the overall result

4. Hints

Some sample questions that you may be asked by tutors in the Weka Laboratory:

i) How do you convert our dataset from *.names/*.data to *.arff format?
ii) How many attributes/records are there in the dataset?
iii) How do you handle the missing values if there are?
iv) What are the classifiers you have built? What are they?
v) Please interpret the classification model.
vi) Which classifier has the best performance? Why?
vii) Which classifier has the worst performance? Why?
viii) Why can't you perform X on our datasets? (where X is one of the classifiers)
ix) …