CSCI3230 Fundamentals of Artificial Intelligence
Weka Lab Specification
Fall 2020

1. Introduction

You are required to use the data mining tool, Weka, with its embedded algorithms and its limitations to work on a provided dataset. Detail is included in the tutorial. There are no laboratory sessions this term and submission is not required.

2. Tasks

i) Download the (installation free) Weka program from our course homepage or the Weka 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. Questions

Choose **three** questions from the list and give your answers.

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)