

## Homework Assignment 2

(100 points)

Assigned Date: Tuesday, February 19, 2013

Due Date: ~~4:00PM Tuesday, February 26, 2013~~  
4:00PM Thursday, February 28, 2013

### Educational Goal

Become familiar with WEKA Java APIs.

### Requirements

- **WEKA data (\Weka-3-6\data):**  
Iris.arff, breast-cancer.arff, soybean.arff, ionosphere.arff, glass.arff
- **Problems**
  1. **Implement a simpleClassify function using a classifier provided by WEKA APIs.**  
**Resources:**  
Netbeans IDE Tutorial for using the Weka API  
[http://www.cs.umb.edu/~ding/classes/480\\_697/homework/WekaJavaAPITutorial.pdf](http://www.cs.umb.edu/~ding/classes/480_697/homework/WekaJavaAPITutorial.pdf)  
Sample code  
[http://www.cs.umb.edu/~ding/classes/480\\_697/homework/WekaTest.java](http://www.cs.umb.edu/~ding/classes/480_697/homework/WekaTest.java)
  2. **Use this simpleClassify function to classify the 5 WEKA data sets.**
  3. **Implement a simpleClassify2 function, which can beat any of the following 5 baseline algorithms: Decision stump, OneR, Decision table, C4.5, PART.**  
(Hint: ensemble learning: a weighted combination of different classifiers may achieve a better performance)
  4. **Write a report to explain the design idea of the simpleClassify2 function including a flowchart and the proposed algorithm in Pseudocode. The report should include detailed experimental analysis.**
  5. **Prepare a readme file to explain how to run the classifiers.**

**Submission Requirements**

1. Submit the softcopy of your report, readme file, and Java source code via UMassOnline. Zero points for late submission.
2. Turn in the paper copy including the cover page of your report, readme file, and Java source code in class. Paper copy should be bound firmly together as one pack (for example, staple, but not limited to, at the left corner). 5 points will be deducted for unbounded homework.
3. No hard copies or soft copies results in 0 points.