

Assignment 2

CIS 472/572 Machine Learning, Winter 2014

due 11:59 pm, Friday, February 7th

1. Write a program to calculate information gains and build a decision tree for the Table 2.1 with the ideas of ID3. Write down your results in the form of IF-THEN rules (i.e., you don't need to draw a tree). The target attribute is EnjoySport (Yes/No).

2. Exercise 3.2 in the textbook.

3. Exercise 3.3 in the textbook.

4. What's the difference between Bagging and Boosting? Based on your understanding, why they can improve the accuracy of classification? What's the main idea of Quinlan to improve the performance of Boosted C4.5?

5. Exercise 4.1 in the textbook.

6. Exercise 4.3 in the textbook.

7. Exercise 4.7 in the textbook.

To turn in by paper version: Put your answers into the homework box for CIS 472/572 in the Deschutes hallway (first floor).

To turn in by emails: If you are in **CIS 472**, email your answers to **hang@cs.uoregon.edu**. If you are in **CIS 572**, email to **dou@cs.uoregon.edu**. We prefer that you send in a pdf file. If you are using Word, you should be able to convert your word file to a pdf file.