

CSCI 6350 Spring 2006**Homework 6**

(Due Thursday, February 23rd)

Name_____

Many research work has been done surrounding “Decision Tree Classification”. This homework requires that you research one of the topics listed below further.

Find at least one research papers (conference or journal paper) on the topic. Write a 1 page summary (12 point, single line) on your findings, with reference to the paper(s). In addition, you need to present your findings to the class on Thursday Feb 23rd. Each person will have 6-7 minutes presentation time. You should prepare for around 10 slides.

Topics:

- Other pruning approaches developed to prevent over-fitting in decision tree
 - Reduced error pruning
 - Pessimistic error pruning
 - Minimum error pruning
 - Critical value pruning
 - Cost-complexity pruning
 - Classifiability based pruning
 - others
- Different approaches developed to handle missing values in decision tree building process
 - C4.5’s approach
 - “Missing is useful” approach
 - other
- Other criteria for selecting best attribute
 - Gini index
 - Gain ratio
 - CART’s approach
 - others
- Feature selection for decision tree
- Regression tree
- Successful real world application of decision tree classification