CSCI 4900/6900 Data Mining

Assignment Number 3: Due 2/27/2014 (in class)

1. [**25 points**] **Short answers please**

   (a) Give a major difference between divide and conquer learning of decision trees and separate and conquer covering methods in data mining?

   (b) Give one advantage to using lift instead of confidence in mining association rules from Market-Basket data.

   (c) Why does a problem occur in Naive Bayes if a particular attribute value does not occur in the training set in conjunction with every class value? Briefly describe a way to fix this problem.

   (d) What is the difference between decision lists and general rule sets?

2. [**50 points**] Do exercise 17.3 on page 571 of the text book.