**MATH 3210**

**Assignment #4 - The Sommelier Project Revisited**

**Scenario:** The dataset Wine.xls contains the values of thirteen attributes for three classes of wines. This data is the result of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars. The three classes of wine do not occur in equal portions and result from a combination of the grape cultivar, the growing conditions and the seasonal weather. In Assignment #2 you analyzed this dataset using descriptive statistics and recommended a rule set that could be used by an automated chemical analysis device to classify the wine type.

**Decision Tree Comparison:** In this assignment you will construct a classification decision tree for the data using C4.5 and compare that result to the result you developed in Assignment #2.

This assignment requires the following operational tasks:

* Download and verify the C4.5 executable
* Download the Wine.xls dataset
* Create a training dataset and a test dataset
* Format the datasets to conform to the application input file structure for the C4.5 application
* Ensure that you save your execution output for analysis
* Compare and contrast these decision tree results with the results of your analysis in Assignment #2

**Write a COMPLETE Analysis Report describing your analysis and recommendations.**