**Decision Trees Lab**

**PRACTICE 1 INSTRUCTION**

Play with different parameters of decision tree operator to generate an easy to understand tree with acceptable level of sophistication that demonstrate factors affecting wine quality. You should use your own judgment as to what tree structure is easy to understand for your manager and yet not too easy that omits important factors affecting wine quality. You should use the training data to predict the wine quality for the scoring data. You manager is only interested in wines of good and excellent quality; you should prepare a list of possible options for him to look at.

It is always a good practice to write an executive summary of your findings and actionable suggestions. In the executive summary you should state the goal of the analysis and summery of your findings and your suggestions for the management. The executive summary should not be more than a paragraph long. It should tell the reader whatever he or she needs to make a decisions without specifying any unnecessary technical or modeling detail.

**DATASET**

*Dataset*: WineQuality-Scoring.csv & WineQuality-Training.csv

*Background*: you have two datasets containing various factors affecting red wine quality. One of the datasets contains the red wine quality attribute and the other does not. The quality attribute is a categorical one with values of excellent, medium, good and mehh.