4/12/2017 Assignment 1

## **Assignment 1**

**Due** Sunday by 11:59pm **Points** 100 **Submitting** a text entry box or a file upload

Available Apr 5 at 12am - Apr 16 at 11:59pm 12 days



## **Low-cost Sensor**

Go to Modules

Your task for this assignment: Design a simple, low-cost sensor that can distinguish between red wine and white wine.

Your sensor must correctly distinguish between red and white wine for at least 95% of the samples in a set of 6497 test samples of red and white wine.

Your technology is capable of sensing the following wine attributes:

- Fixed acidity Free sulphur dioxide
- · Volatile acidity Total sulphur dioxide
- Citric acid Sulphates
- Residual sugar pH
- · Chlorides Alcohol
- Density

## **Tasks**

- 1. Read WineQuality.pdf.
- 2. Use the RedWhiteWine.csv/arff file that is provided.

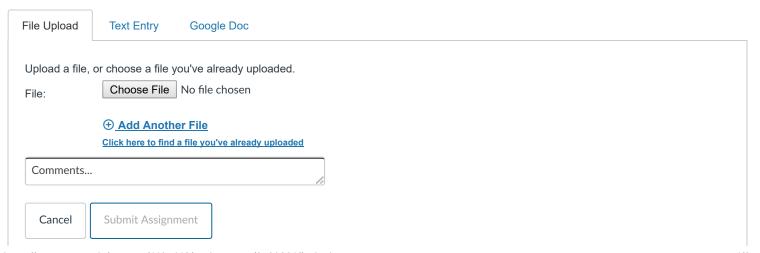
Note: If needed, remove the quality attribute, which you will not need for this assignment.

- 3. Build an experiment using Decision Trees.
  - 1. What is the percentage of correct classification results (using all attributes)?
  - 2. What is the percentage of correct classification results (using a subset of the attributes)?
  - 3. What is the AUC of your model?
  - 4. Visualize your decision tree
- 4. What is the best AUC that you can achieve?
- 5. Which are the the minimum number of attributes? Why?

## Submission

Submit a 1-pager highlight your solution, and answering the 5 questions above

- 1. You can use either R or Python.
  - Ideally, share the link for your online notebooks, so we can reproduce your results.
- 2. Note your findings and any insights when working on this assignment



4/12/2017 Assignment 1