CS 5751 – Spring 2018 – Homework 7
Assigned: 03/27/2018
Due: 04/03/2018
Total points: 100 pts.
Submit a soft copy to canvas. Remember to write your name at the top of each file you submit.

Objectives: The objectives of this homework are the following:

- Learn how to fit kNN models.

Notes:

- This homework is to be done individually. You may discuss with your classmates, but the work that you write must be your own.

Activity: (100 pts.) (Classification models) Using either R or Python, do the following:

a) (0 pts.) Download the wine dataset from the UCI website.
b) (20 pts.) Do any pre-processing that may be needed.
c) (30 pts.) Build kNN models with K=1,2,3,…,10 to classify the quality of the wine. Remember that for these models you need to normalize the data.
d) (30 pts.) Build a decision tree model to classify the quality of the wine. Add code to "pretty print" your tree (if it is too big to display, then do your best).
e) (20 pts.) Compute the error rate for each model using K-fold cross validation, for K=10. Comment on the results you obtain.

For this activity, write a Jupyter notebook named yourLastName_hw7.ipynb that implements 1a through 1e.