HW #3: MATLAB/Programming

In this homework, you will practice importing data from a text file, applying matrix manipulations, and using functions/libraries.

Download the WINE dataset from: http://archive.ics.uci.edu/ml/datasets/Wine

1. Import the data into MATLAB. (see doc csvread)
2. Split the data into a 50/50 training set and a testing set. You should train with 50% of the samples from each class. Note that they are unbalanced (unequal number samples).
3. Use linear SVM (support vector machine) to predict the class of the testing set. What is the overall accuracy? How about the accuracy for predicting each of the classes? Can you do this without calling the prediction function several times?
4. (Optional) Can you get the leave one out cross validation accuracy given all data? Note that an efficient implementation is already done for LIBSVM.

Notes:

- You may need to paste the WINE data into a .txt file. Make sure you use a *text editor* like *notepad* to keep the data clean. Programs like Word add extra stuff to txt files.
- The class labels are the first column of the data
- 'find' is a useful function for returning the indices of a particular class.
- You can combine a matrix A and B in a new one by: C = [A B]
- You can index rows: C = A([1:10], :), columns: C = A(:, [1:10])
- You can remove the first two columns of a matrix by: C(:, 1:2) = []

- Given the data, think about your approach before blindly coding. Check that your data is in the appropriate form before calling the SVM functions.