v.2

Assignment – 1 (5+1 bonus pts)
EECS 4404/5327

The assignment is due on **Monday, October 14, 2019** by 6p.m..

Submit a .zip package of your work including a single pdf file of your assignment with your solutions, each question at a new page, plus a folder containing your Matlab code, each question as a separate .m file, on Moodle's respective assignment tab.

Note. The goal of this assignment is for you to learn these concepts in practice. You are not allowed to use MATLAB's built-in functions for fitting curves. However, if you exhausted your effort and did not manage to come up with a solution, using those functions will get you **2/5** points on this assignment.

We use wine dataset already available in Matlab. It can be accessed by

```
[x,t] = wine_dataset;
```

Alternatively, it can be downloaded from
http://www.mediafire.com/file/dfmmwxumxfh3ifv/wine.mat/file

It contains 178 different wines (observations) from 3 winery (labels) with these 13 features:

1. Alcohol
2. Malic acid
3. Ash
4. Alkalinity of ash
5. Magnesium
6. Total phenols
7. Flavonoids
8. Nonflavonoid phenols
9. Proanthocyanidins
10. Color intensity
11. Hue
12. OD280/OD315 of diluted wines
13. Proline

The last column of the wine.mat file(if downloaded), or, the variable t (if you use Matlab's built-in data) has the labels of each wine, meaning that it belongs to one of the three wineries.

v.2

Question-0 (Preprocessing)

Remove all row corresponding to the labeled winery 3. After this process, you should have only 2 labels on your data.

Question-1 (0.25 pts)

Load the data and plot (visualize) the data points of wines by their Alcohol (feature 1 in x axis) and Malic acid (feature 2 in y axis).

Question-2 (1 pts)

Pick Magnesium and Color intensity as your two features and for degrees n =1, ..., 10 fit a polynomial of degree n to your data. Plot those fitting lines on the data. You can check the correctness of your solution with MALAB's built-in curve fitting function.

Question-3 (1 pts)

For each learned function (n=1, ..., 10), compute the empirical square loss (ERM) on data and plot it as a function of n.

Question-4 (1 pts)

Now, fix the n=10 and add a lasso regularization for your predictor of data. Vary the regularization parameter in a loop of 20 and visualize the RLM loss. You can check the correctness of your solution with MALAB's built-in Lasso.

Question-5 (0.25 pts)

Now, add a third feature of Hue to your data and plot the three in a 3D plot.

Question-6 (1 pts)

For your three selected features, fit a surface to your data of a degree 10.

Question-7 (0.5 pts)

Compare the ERM loss of your surface (question 6) and line (question 3) predictors.

Question-8 (1 bonus pts)

Fit the data with a Perceptron classifier and compare the loss with respect to your fitted lines (question-3)