CS 5785 COMBINED-XLIST

Applied Machine Learning

HW 0

Wentao Xu (wx225)
Xingyue Dai (xd86)

General summary:

Following the instructions, we have successfully set up the working environment for this course. Specifically we have installed Anaconda that includes Python and jupyter notebook which we used for this homework.
We downloaded Iris Flowers dataset from the internet, and as you might see from the .ipynb file under the same folder, we parsed the data and put it into an 151 x 5 array as there are 151 samples with 5 atrributes.

```
['5.1', '3.5', '1.4', '0.2', 'Iris-setosa']
['4.9', '3.0', '1.4', '0.2', 'Iris-setosa']
['4.7', '3.2', '1.3', '0.2', 'Iris-setosa']
['4.6', '3.1', '1.5', '0.2', 'Iris-setosa']
```

To visualize the dataset, we applied matplotlib to draw a scatterplot as shown below (see source code in the .ipynb file)