2017 Fall: COMP-SCI 5590/490 - Special Topics

Python Programming

Lab Assignment 3

**Assignment Overview**

The following assignment focus on to make one familiar with python machine learning

**Lab Assignment**

1. Pick any dataset from the dataset sheet in class sheet and make one prediction model using your imagination with Linear regression. Some examples are:

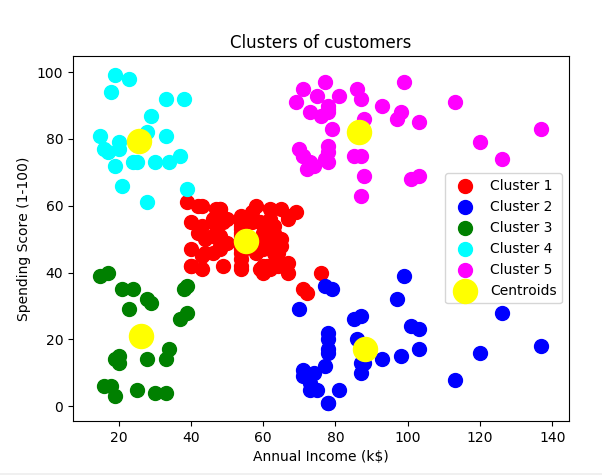
* Predicting weather
* Predicting salary according to age

You can also pick dataset of your own.

1. Implement Kmeans clustering on customers.csv file provided into five clusters.

Dataset link: <https://umkc.box.com/s/ksnolajzn2gzo1723grj4hztm7rc9pu9>

Sample output:



1. Implement SVM classification,
2. Choose one of the dataset using the datasets features in the scikit-learn(you should not select the same dataset in the class)
3. Load the dataset
4. According to your dataset, split the data to 20% testing data, 80% training data(you can also use any other number)
5. Fit the model on the training data
6. First apply linear kernel
7. Apply the accuracy
8. Then apply RBF kernel instead of linear
9. Report your view if changing the kernel affect the accuracy of your model
10. Write a program

Take Input of the project is one text file. Use the simple approach below to summarize a text file:

* Read the file
* Using Lemmatization, apply lemmatization on the remaining words
* Using POS, remove all the verbs
* Calculate the word frequency of the remaining words
* Choose top five words that has been repeated most
* Go through the original text that you had in the file
* Find all the sentences with those most repeated words
* Extract those sentences and concatenate
* Enjoy the summarization

**Submission Guidelines:**

* Submit your code at Github and properly document it. Submit your screenshots as well.
* Properly document your code
* Submit only the code portion in text file to UMKC blackboard assignment.
* Remember code similarity to be less than 45%
* Use following link to submit your assignment:

<https://goo.gl/forms/cxvY8Kg1pvNNzrpw1>