

Mail Classifier
Programming Club Summer Project, 2017

Arnav Garg (160147)
Email ID : arnavgar@iitk.ac.in

Kshitiz Tyagi(160348)
Email ID : ktyagi@iitk.ac.in

July 7, 2017

1 Week 1

1. Started learning python, right from naive concepts of variables, operator to loops, functions while learning the syntax used.
2. Delved into slightly advanced topics like classes, getting brushed up with Object Oriented Programming(OOP).

2 Week 2

1. Started Udacity course on Machine-learning, learning how it works practically.
2. Learned about Naive-bayes Algorithm, and how to apply it on a given dataset consisting of features and labels.
3. Learned about finding accuracy of the classifier using different modules.

3 Week 3

1. Learned about some other new algorithms like SVM, Tree classifier.
2. Applied the concepts to some actual dataset comprising of emails of two persons and then predicting which email belongs to whom. Learned how the tree-classifier works on a recursive approach.
3. Tried to increase accuracy of the classifier by changing different parameters.

4 Week 4

1. Continued the course on udacity. Looked into some other algorithms like 'K-Nearest-Neighbours'. Learned about datasets, particularly Enron-dataset and how are we supposed to deal with it.
2. Learned concepts of regression, 'outliers' and where and how they are used.

5 Week 5

1. Learned about 'text-processing' and 'stemming' to know how the data would be used by the code.
2. Tried parsing out the data and pre-processing it. Overall, got the basic understood the structure of how the classifying code works.

6 Week 6

1. Started writing the code and finished it.
2. Debugged the code , and tried to improve its accuracy by increasing its test data.