Data Mining

Assignment 1

Due date: 4ᵗʰ March 2019

1. The assignment deals with the implementation of K-NN classifier.
2. The source data is the IRIS data set which is available with the **scikit-learn** package
3. You are expected to code the assignment using **Python**.
4. The scikit-learn package provides toy datasets such as the IRIS one.
5. The IRIS dataset seeks to classify different plants based on their appearance features.
6. Import the data from the **sklearn.datasets** and divide into test and training samples.
7. Code your own K-NN classifier and run on the data.
8. Plot accuracy for different values of K (say from 1 to 10)
9. Normalize and standardize your data and observe if that makes your classifier achieve higher accuracy.
10. Please provide the code file as well as the plots generated against K values as the deliverables.