



California State University Long Beach
BME 455 Spring 2021
Lab 10
Instructor: Elsa Harris

In this lab you will work with Clustering algorithms

Objectives:

1. Perform clustering from scratch on a simple dataset to understand how the algorithm works
2. Perform k-means Clustering using the scikit-learn
3. Calculate the Silhouette Coefficient to assess the performance of the clustering algorithm
4. Predict which cluster a new unseen point belongs to
5. Find the optimal k for a k-means algorithm

Part 1: Follow the instructions on Chapter 10 of “Python Machine Learning” by Wei-Meng Lee.

Part 2: Follow the instructions on Chapter 4 - “Data Science with Python” by Oreilly from page3 to page 6 on the pdf for HCA section.

Part3: Prepare a one page report on what you learned about the k-means clustering algorithm. Make sure to list the main python and scikit-learn commands to perform the tasks 1-5 as listed above under “Objectives”

What to submit:

- a. A pdf printout of your complete code. Please run the code and make sure the outputs (numerical values, graphs, etc) are printed in the file you submit.
- b. A Python notebook file
- c. A pdf printout of your report as in Part 2 above

Name each submission file as Lab10_PartX_FirstName_LastName