Assignment #2 - Unsupervised Learning

Firas Bayram
firas.bayram@kau.se

Karlstad University — August 26, 2021

Introduction

Unsupervised machine learning algorithms are those algorithms where only the input dataset is available, not the corresponding output. The learning task is to find the function that discovers hidden patterns and the structure of our input dataset. This assignment aims to get you familiar with unsupervised machine learning algorithms and understanding how it works in practice. The goal of the assignment is to give hands-on exposure to the machine learning algorithms we have discussed in class: K-means and hierarchical clustering. In this assignment, we will use the Python programming language and Jupyter Notebook to develop the solution.

Part # – K-Means

1. Generate and plot a 2D dataset with c classes of 1000 points.
2. Select k the number of clusters equal to the c in Step 1, implement K-means algorithm with cluster centers initialized both randomly and using K-means++ algorithm.
3. Visualize the cluster centers and the results of Step 2.
4. Repeat the previous steps by running the code multiple times.

Part #2 – Hierarchical Clustering

1. Generate and plot a 2D dataset with c classes of 1000 points.
2. Select k the number of clusters equal to the c in Step 1, implement Agglomerative Hierarchical Clustering algorithm.
3. Visualize the results of Step 2 and the corresponding dendogram of the clustering algorithm.
4. Repeat the previous steps by selecting different c in Step 1.

Examination

Hand in your solution in Jupyter Notebook on Canvas and book a meeting to demonstrate your solution with one of the lab assistants. Note that lab demonstration meetings are booked through email, not on Slack.

End of Assignment