

# Lab 3: Clustering of Iris Flowers.

---

During this lab you will gain practical knowledge of the clustering methods that have been presented to use during lectures 18-20. To refresh the material, please consult the notes available for the module on Moodle.

In this lab, you will need to use **SciKitLearn** module of Python to import contents of iris dataset. Assume that the dataset provided by **SciKitLearn** is clean and free of errors. Assuming that the original Isi classification is unknown, use the techniques presented to you in class to determine the cluster-arrangement of the data points for a variety of preset cluster numbers/configurations. Which cluster number represents the natural cluster arrangement of the data points.

## Classification of Iris Flowers (10 points)

**Step 1.** Import the Iris Dataset from SciKitLearn. **(1 point)**

### K-Means clustering (6 points)

**Step 2.** Use K-Means to build 2, 3, 4, ... 10 clusters. **(1 point)**

**Step 3.** Plot values of the within cluster distance with respect to the number of clusters. **(1 point)**

**Step 4.** Plot values of the between cluster distance with respect to the number of clusters. **(1 point)**

**Step 5.** Plot values of the Calinski-Herbasz index with respect to the number of clusters. **(1 point)**

**Step 6.** What is the natural cluster arrangement and why? **(2 point)**

### Hierarchical Clustering (3 points)

**Step 7.** Use Hierarchical clustering to identify arrangement of the data-points. **(1 point)**

**Step 8.** What is the natural arrangement there and why? **(2 point)**