Machine Learning

Session 5 - PL

Unsupervised Learning - Clustering

Ciência de Dados Aplicada

2023/2024

Clustering in Python (scikit-learn + scipy)

- https://scikit-learn.org/stable/modules/clustering.html

- In scikit-learn clustrering mehods follow the fit-predict methods:
  - "fit" compute the clusters;
  - "predict" predicts the cluster of the samples;

- https://docs.scipy.org/doc/scipy/reference/cluster.html

Clustering in Python (scikit-learn + scipy)

- Scikit-learn:
  - K-means:
    - https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html
  - Hierarchical:
    - https://scikit-learn.org/stable/modules/generated/sklearn.cluster.AgglomerativeClustering.html

- Scipy:
  - K-means:
    - https://docs.scipy.org/doc/scipy/reference/generated/scipy.cluster.vq.kmeans.html
  - Hierarchical:
    - https://docs.scipy.org/doc/scipy/reference/generated/scipy.cluster.hierarchy.linkage.html

Exercises:

- Notebooks on the github repository:
  - Notebook with examples:
    - notebooks/session5/examples.ipynb
  - Notebook with exercises:
    - notebooks/session5/exercises.ipynb