← Go Back to Data Analysis & Visualization

Course Content

New Package Introduction - Unsupervised Learning: Scikit-Learn-Extra

Scikit-Learn-Extra

The scikit-learn-extra is a Python module for machine learning that extends scikit-learn. It includes algorithms that are useful but do not satisfy the scikit-learn inclusion criteria, for instance, due to their novelty or lower citation number.

Library installation

The latest release can be installed with conda, by running the below command in **Anaconda Prompt (or terminal for Mac)**.

```
conda install -c conda-forge scikit-learn-extra
```

Alternatively, we can install using the pip command

```
pip install scikit-learn-extra
```

User Guide

This library helps us to implement the K-Medoids algorithm on the data. K-Medoids can be more robust to noise and outliers in comparison to K-Means as it will choose one of the cluster members as the medoid while K-Means uses the mean which can get affected by the outliers in the data.

To import the K-Medoids function,

```
from sklearn_extra.cluster import KMedoids
model = KMedoids()
```

Parameters:

metric: distance metric to be used (default: 'euclidean')

n_clusters: number of clusters to be formed and hence the number of medoids (one per cluster) (default value: 8)

init: 'heuristic' method used for medoid initialization

max_iter: maximum number of the algorithm's iterations to be performed when fitting the data.

Example

```
from sklearn_extra.cluster import KMedoids

kmedo = KMedoids(n_clusters = 3, random_state = 1)

kmedo.fit(data_scaled)
```

In the above code, we are implementing the K-Medoids clustering algorithm with 3 clusters having random state 1.

For your reference, go through this documentation page of the Scikit-Learn-Extra package.

Previous

Next

Help