K-Shape (tslearn) for S1/S2 characterization

Emmanuel Le Guirriec (CPPM)

DarkSide CPPM meeting

14 December 2021

Simu 100 events 1 photon 100keV with noise

- **Extract segments where mountains found**
  - Create a pickle file with 225 segments
  - If segments start before 4000, S1
    - Else S2

	segment	event	mount	start
0	[0, 4, 5, 13, 16, 13, 18, 13, 10, 10, 11, 6, 7...	0	0	1249
1	[0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, ...	0	1	26827
2	[0, 2, 4, 6, 20, 15, 16, 14, 10, 10, 14, 12, 9...	1	0	1248
3	[0, 2, 0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...	1	1	23971
4	[0, 2, 2, 0, 0, 0, 2, 0, 0, 6, 8, 12, 22, 21, ...	2	0	1241
..	...	...	...	...
220	[0, 2, 0, 0, 0, 0, 0, 0, 0, 2, 4, 12, 16, 11, ...	99	0	1240
221	[0, 3, 0, 0, 0, 2, 0, 0, 2, 2, 0, 2, 3, 5, 2, ...	99	1	4633
222	[0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...	99	2	7250
223	[0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ...	99	3	11590
224	[0, 2, 0, 0, 0, 0, 0, 0, 0, 0, 3, 3, 0, 0, 0, ...	99	4	24438

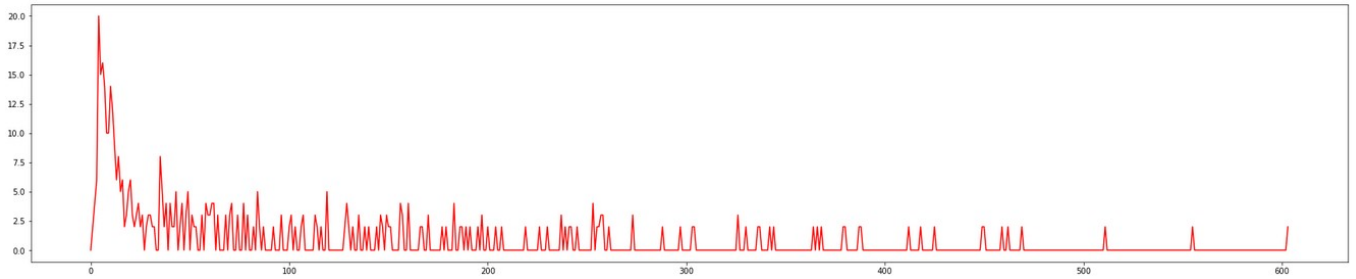
[225 rows x 4 columns]

tslearn package

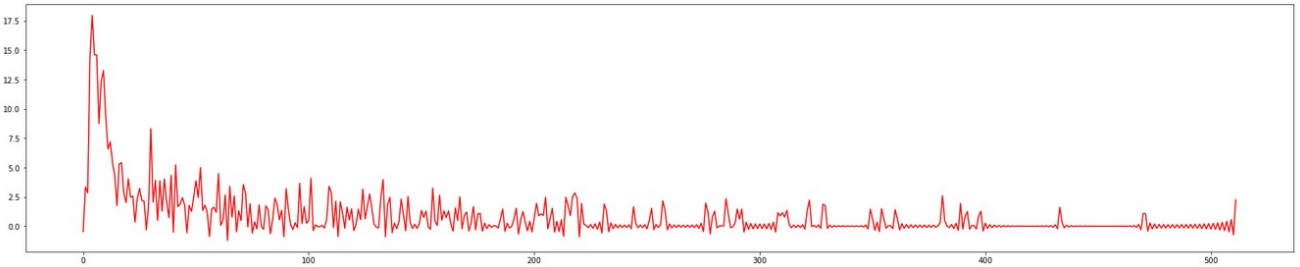
- **tslearn is a Python package that provides machine learning tools for the analysis of time series. This package builds on scikit-learn, numpy and scipy libraries.**
  - https://tslearn.readthedocs.io/en/stable/index.html
- **Time Series Clustering methods of tslearn (unsupervised method)**
  - KernelKMeans (not tested)
  - TimeSeriesKMeans
    - With different metrics (Euclidean, DTW, SoftDTW)
  - k-Shape
    - **k-Shape preserves the shapes of time series computing centroids (average sequence) effectively under the scaling and shift invariances**
    - k number of clusters (2, 3, 4)

Prepare data

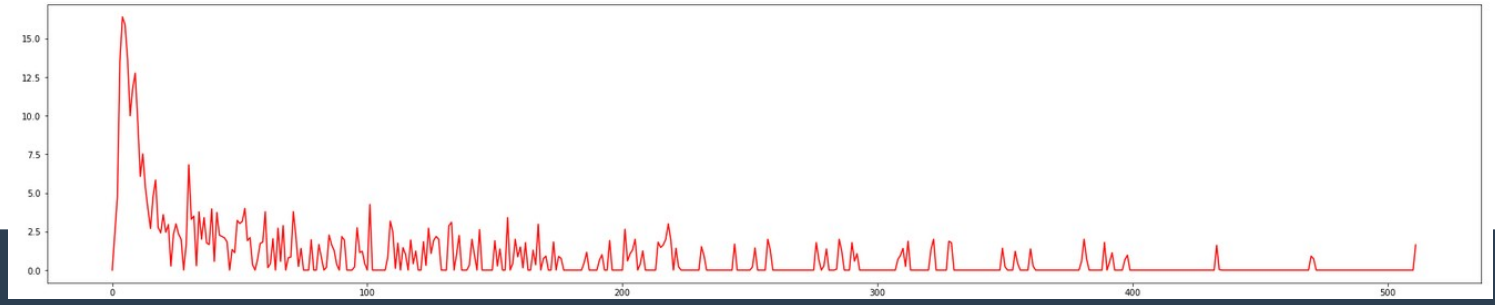
- **Different segment lengths**
  - Some clustering methods need the same size
    - Try with 512



  - resample method of scipy.signal



  - resample by interpolation
    - https://github.com/nwhitehead/swmixer/blob/master/swmixer.py



Results

- k-Shape is the best method
- Requesting 2 clustering classes
  - All the S1 are well characterized
  - 2 S2 are badly characterized