Advisors: Bruno Focassio. Gabriel Schleder &
Yasmin Watanabe

EXPLORING TIME-SERIES MOTIFS THROUGH DTW-SOM

Júlia Guedes & Yasmin Shimizu

TABLE OF CONTENTS

Motif Discovery

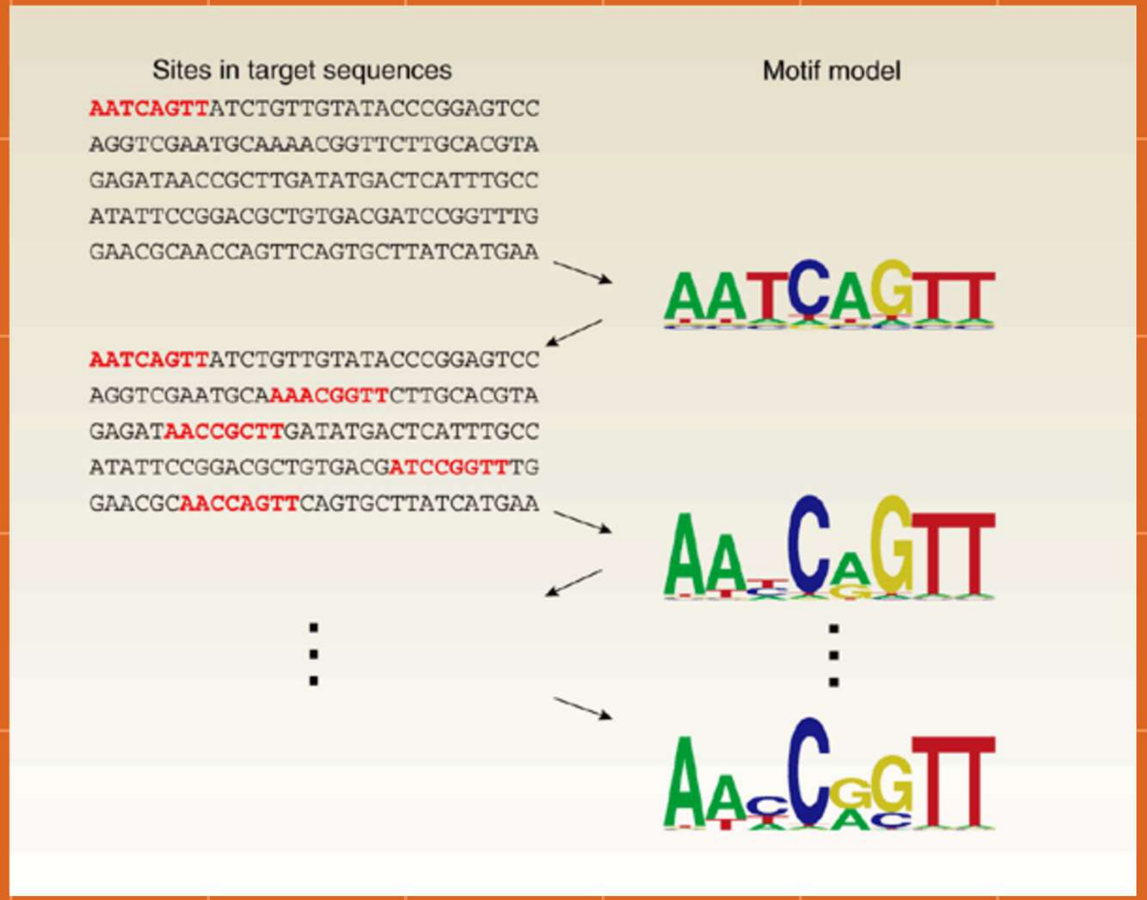
DTW

SOM

DTW-SOM

Examples

How does a DNA sequence motif discovery work?



D'haeseleer, 2006

MOTIF DISCOVERY

Extraction of all (or a specific subset) of the significant segments in a large dataset.

Similarity-based

Focus on. the similarity of the segments.

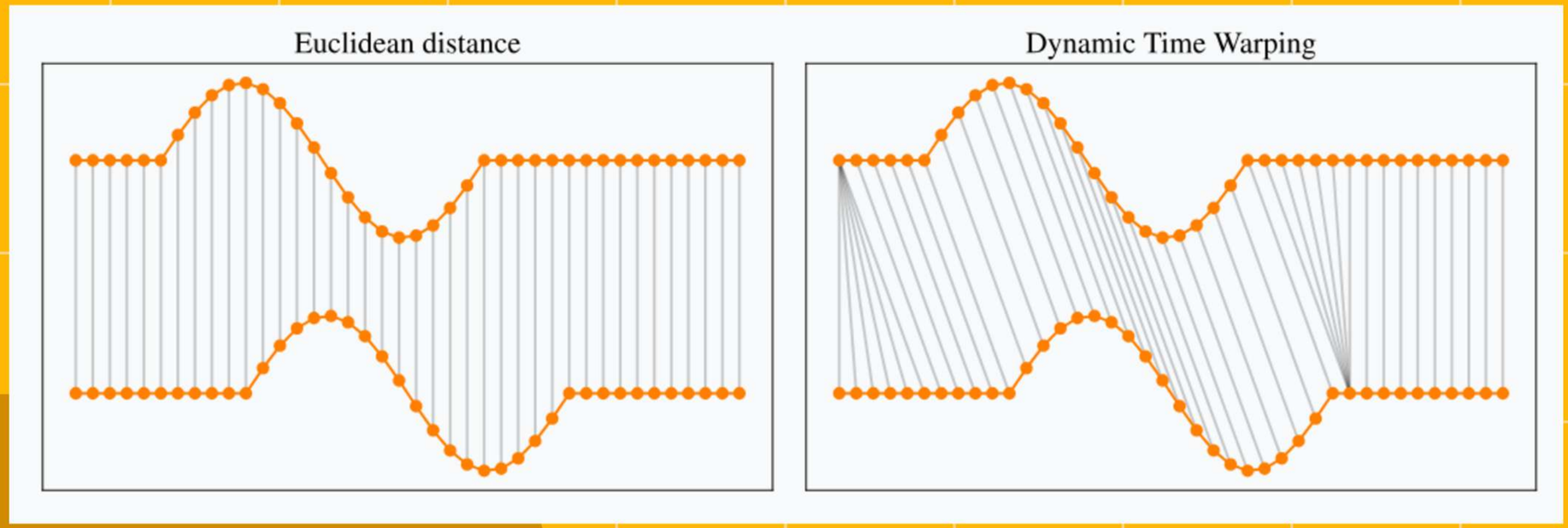
Support-based

Focus on the repetition of the segments.

DYNAMIC TIME WARPING (DTW)

A technique used to compare two time series datasets when the time indices between corresponding data points are not perfectly aligned.

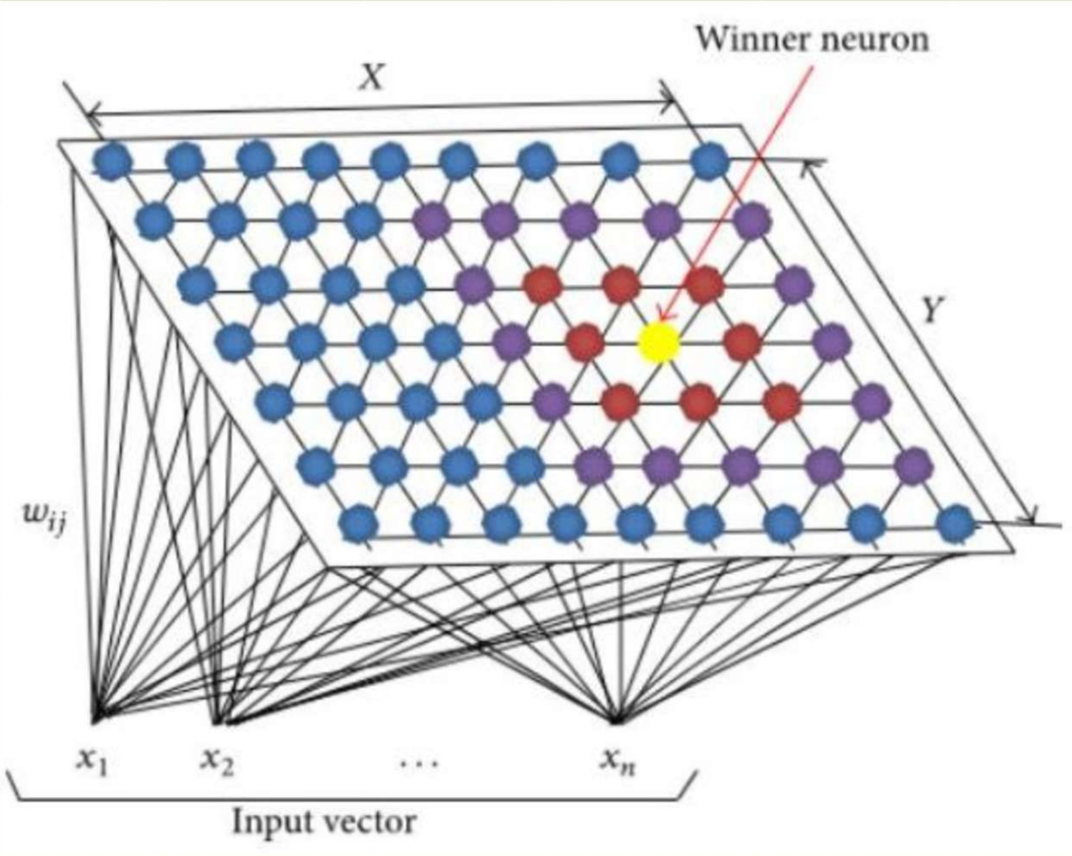
Comparison between DTW and Euclidean distance



Tavenard, 2021

SELF ORGANIZING MAP (SOM)

Architeture of a SOM algorithm



Mayatopani, 2023

An unsupervised artificial neural network algorithm used to map high-dimensional data onto a low-dimensional, discrete feature map, **preserving the relationships** between data patterns.

$$m_i = m_i + \alpha(t)h_{ci}(t)\,|x - m_i|$$

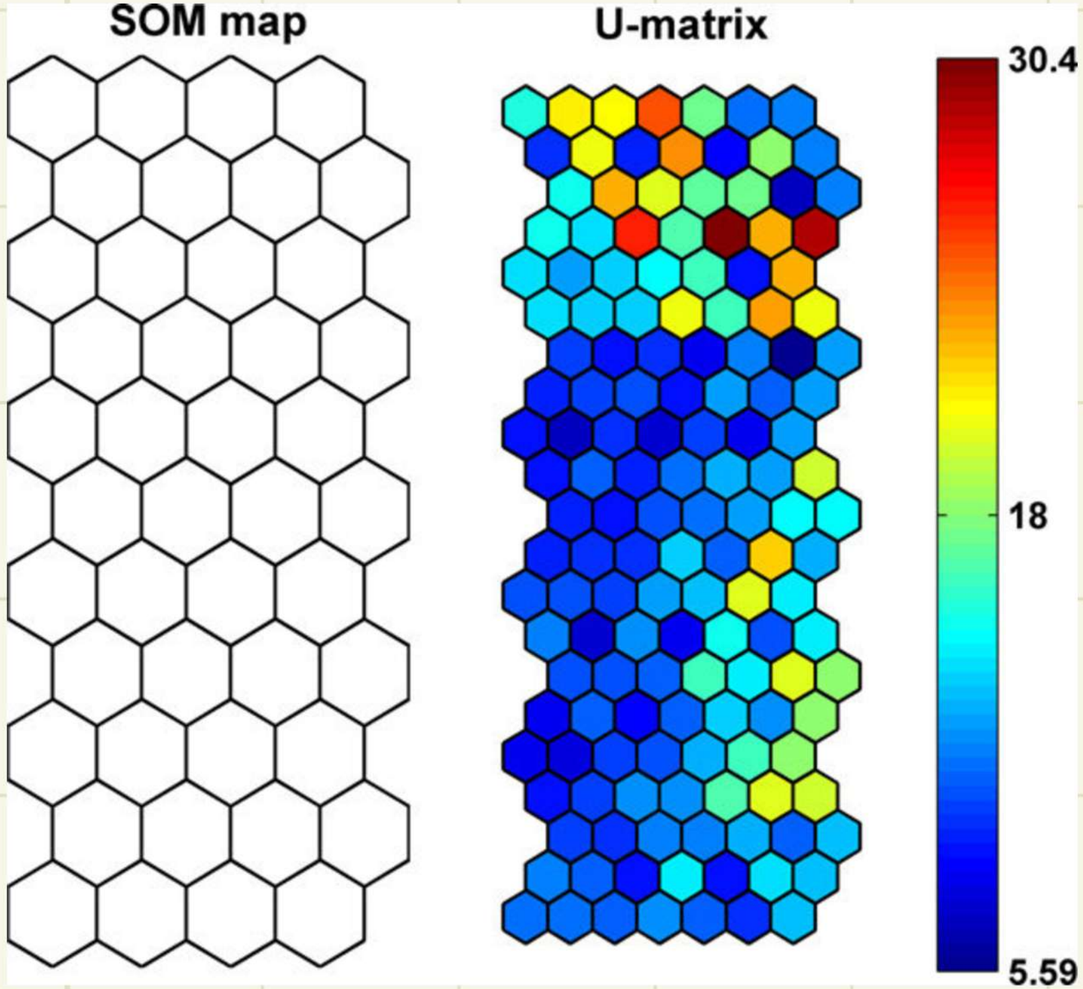
Weights actualization

VISUALIZATION TOOLS:

UNIFIED DISTANCE MATRIX (U-MATRIX)

A visual representation of the SOM that illustrates the average distances between neighboring neurons in the input data space.

Representation of a SOM map and its respective U-matrix



Bieroza, 2012

VISUALIZATION TOOLS:

WINNER MATRIX

A visual representation where each element corresponds to a neuron, and the value represents the number of input data objects it has won during the final phase of training.

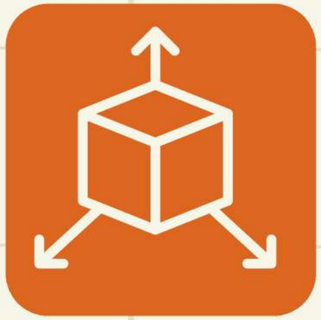
Representation of a Winner Matrix

PyClustering Documentation

DTW-SOM

Compare time series segments

Receive fixed-length and variable-length motifs.



Represent multi-dimensional motifs

consider a given order of significance

Visualization of the relationships between the motifs

EXAMPLES

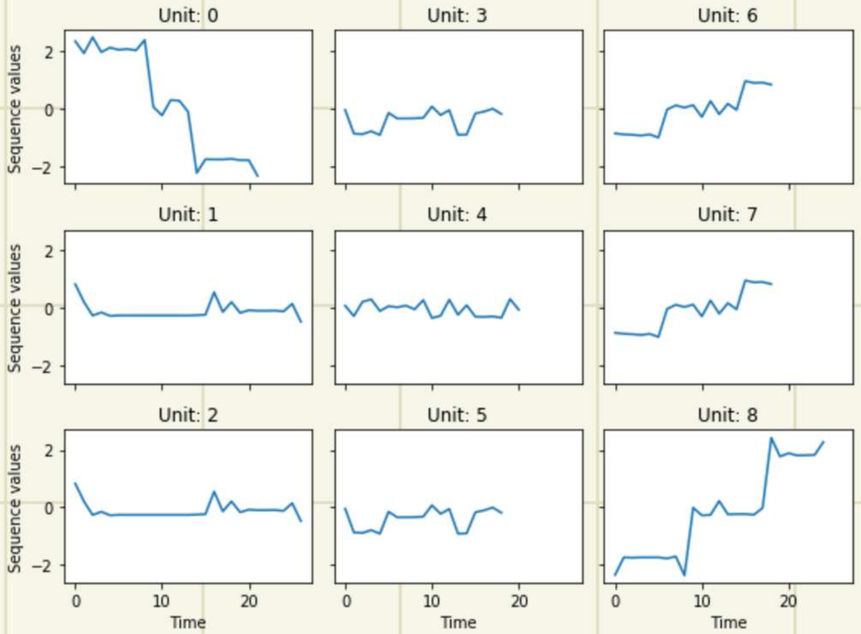
- Experimental with the **synthetic motif** dataset
- Experimente with the **GunPoint** dataset
- Experiment with the **UWaveGesture** dataset

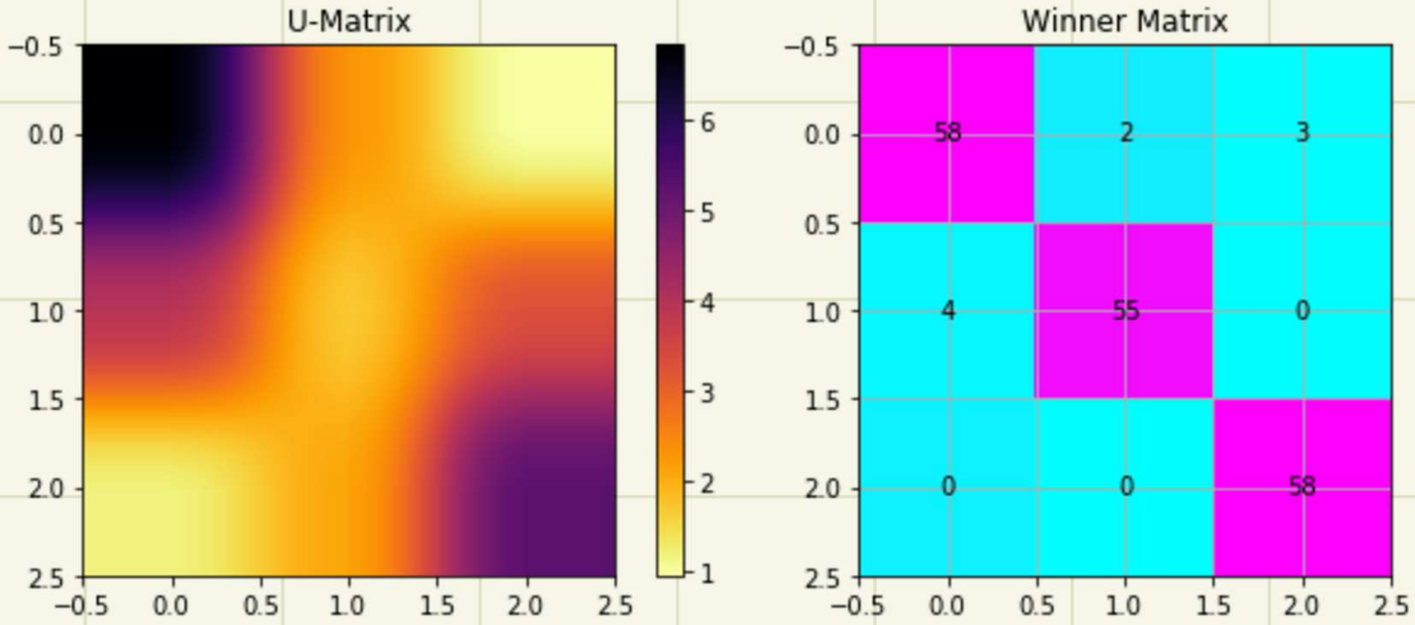
SYNTHETIC MOTIF DATASET

A synthetic dataset of **motifs centers** forming 3 clear clusters to evaluate the efficiency of the method DTW-SOM.
If these clusters were detectable, then the DTW-SOM was working as expected.

RESULTS

Sequence values of the units, U-Matrix and Winner Matrix after training the DTW-SOM with synthetic motifs.





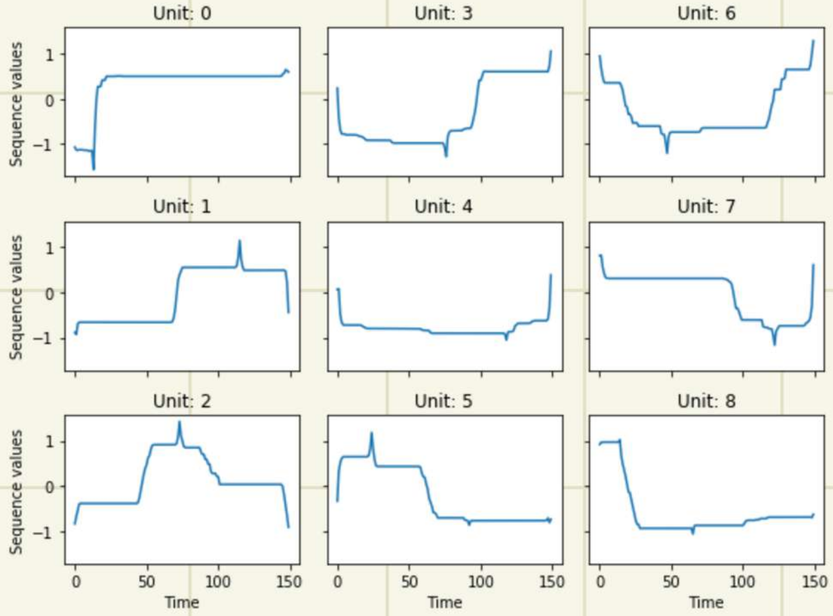
Silva & Henriques, 2020

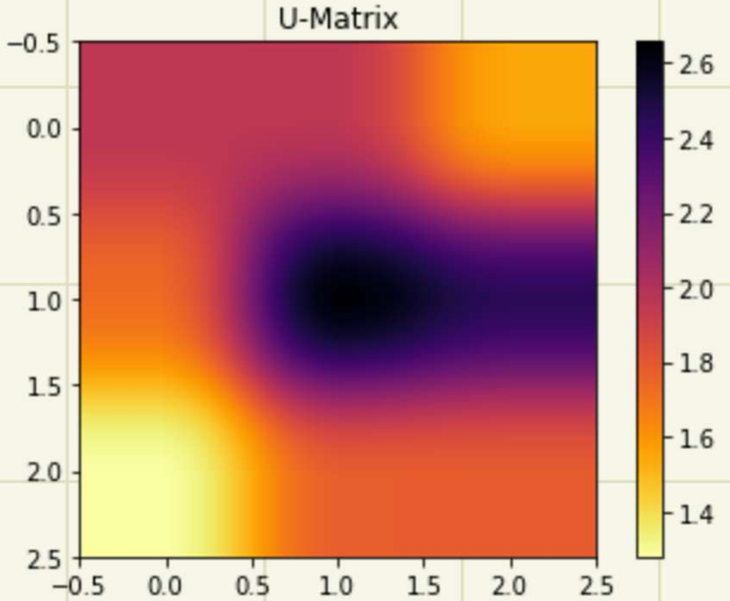
GUNPOINT DATASET

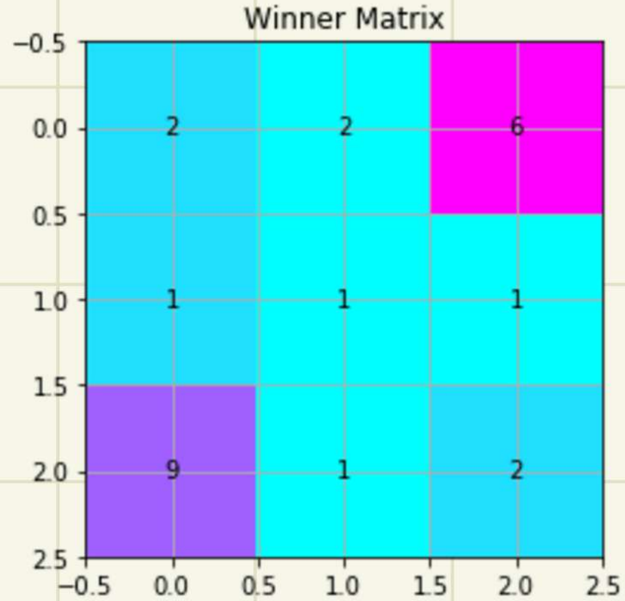
A time-series analysis tracking the centroid of the **hand's motion** along the x-axis, capturing two distinct actions: drawing a gun and pointing a finger.

RESULTS

Sequence values of the units, U-Matrix and Winner Matrix after training the DTW-SOM with GunPoint.







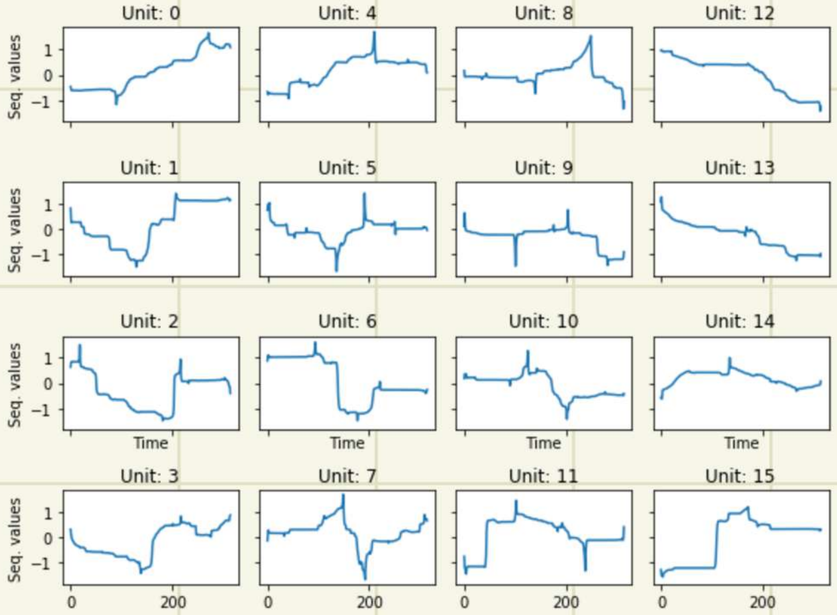
Silva & Henriques, 2020

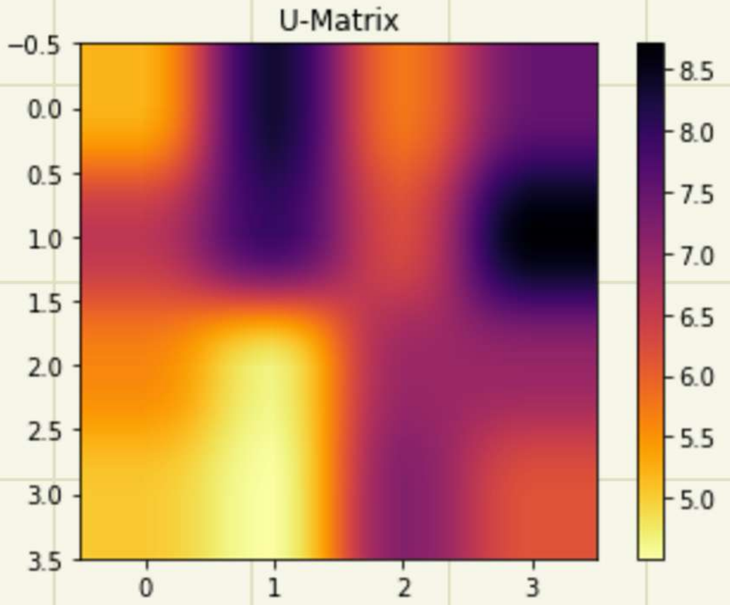
UWAVEGESTURE DATASET

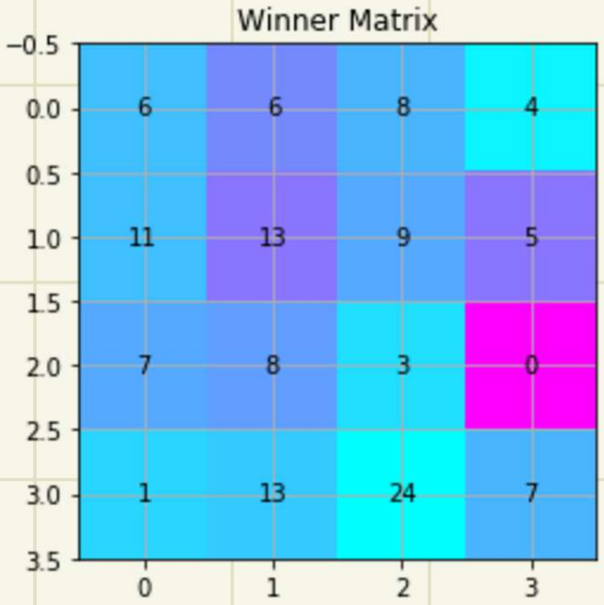
Recording of **acceleration** measurements of 8 especifics **hand gestures** in the space. Here, it's used only the x-axis dataset, about lateral movements.

RESULTS

Sequence values of the units, U-Matrix and Winner Matrix after training the DTW-SOM with UWaveGesture.







Silva & Henriques, 2020

CONCLUSION

- DTW-SOM is capable of extracting relevant information from a set of motifs and display it in a space-efficient way, being useful to identify data patterns and understand results
- During the experiment with the synthetic dataset, they observed that the random sample initialization was not as robust as the anchors initialization.
- As future work, they propose an investigation on more robust initialization schemes.

REFERENCES

[1] Silva, Maria Inês, e Roberto Henriques. Exploring time-series motifs through DTW-SOM. arXiv:2004.08176, arXiv, 17 de abril de 2020. arXiv.org, https://doi.org/10.48550/arXiv.2004.08176.

[2] D'haeseleer, Patrik. How Does DNA Sequence Motif Discovery Work? Nature Biotechnology, vol. 24, no 8, agosto de 2006, p. 959–61. www.nature.com, https://doi.org/10.1038/nbt0806-959.

[3] "Understanding Dynamic Time Warping". Databricks, 30 de abril de 2019, https://www.databricks.com/blog/2019/04/30/understanding-dynamic-time-warping.html.

[4]Tavenard, Romain. An introduction to Dynamic Time Warpi, 2021.. https://rtavenar.github.io/blog/dtw.html. Acesso em 28 de janeiro de 2025.

[5] Mayatopani, Hendra, et al. "Implementation of Self-Organizing Map (SOM) Algorithm for Image Classification of Medicinal Weeds". Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi), vol. 7, no 3, junho de 2023, p. 437–44. jurnal.iaii.or.id, https://doi.org/10.29207/resti.v7i3.4755.

[6] Bieroza, Magdalena & Baker, Andy & Bridgeman, John. (2012). Exploratory analysis of excitation-emission matrix fluorescence spectra with self-organizing maps-A tutorial. Education for Chemical Engineers. 7. 10.1016/j.ece.2011.10.002.

[7] pyclustering.nnet.som.som Class Reference. https://pyclustering.github.io/docs/0.8.2/html/d7/d7b/classpyclustering_1_1nnet_1_1som_1_1som.html. Acesso em 28 de janeiro de 2025.

THANK YOU :)

ANY QUESTIONS?

Júlia Guedes &
Yasmin Shimizu