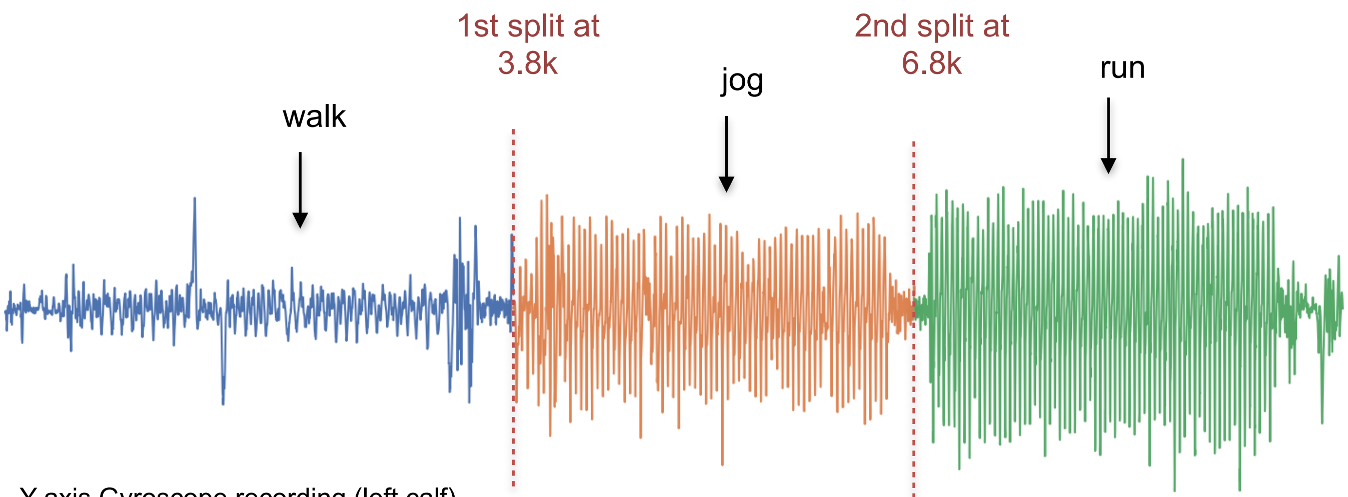
Practical Time Series Classification: Time Series Segmentation

DSAA, 12.10.2025, Birmingham, United Kingdom
Arik Ermshaus

Table of Contents

- **Time Series Segmentation Task & Use Cases**
- Algorithms
- Use Cases & Limitations

Time Series Segmentation (TSS)

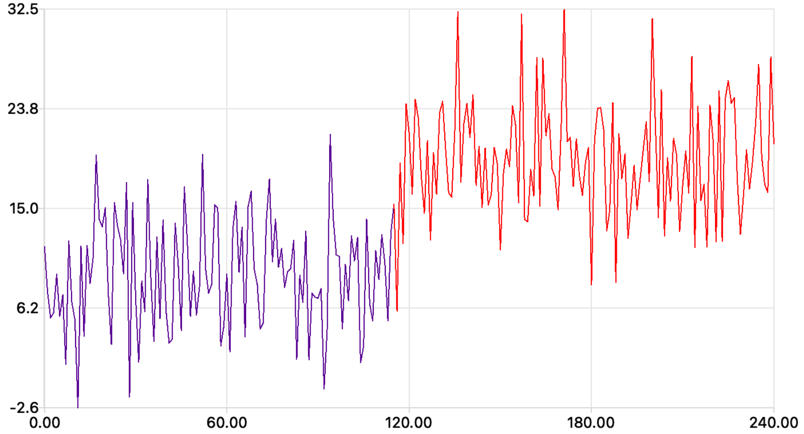


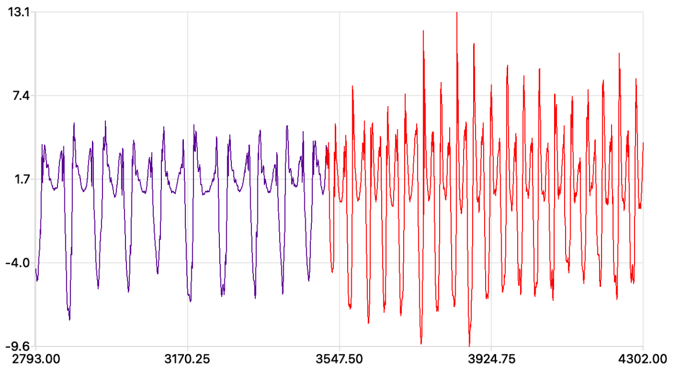
Y-axis Gyroscope recording (left calf)

Banos, O., Toth, M. A., Damas, M. et al. (2014). Dealing with the effects of sensor displacement in wearable activity recognition.

- TSS partitions TS into variable-sized meaningful segments; based on change points
- Notion of “meaningful” depends on domain
- Complex unsupervised preprocessing for TS classification

TSS: Types of Temporal Changes

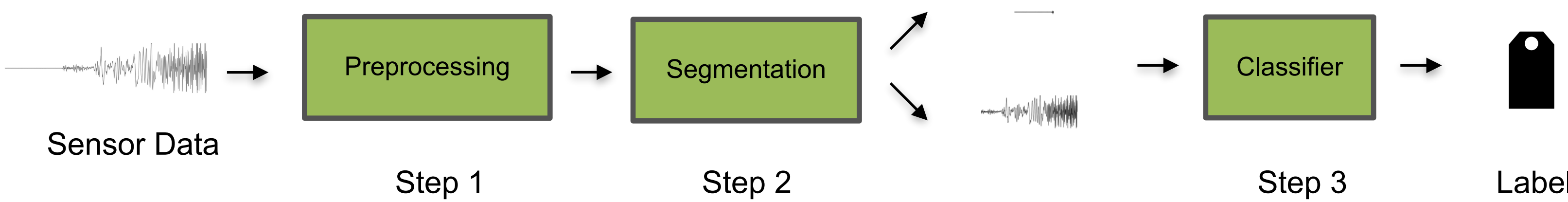




- Changes in distributions
- Piecewise statistics
- e.g. mean, var, trend, sine

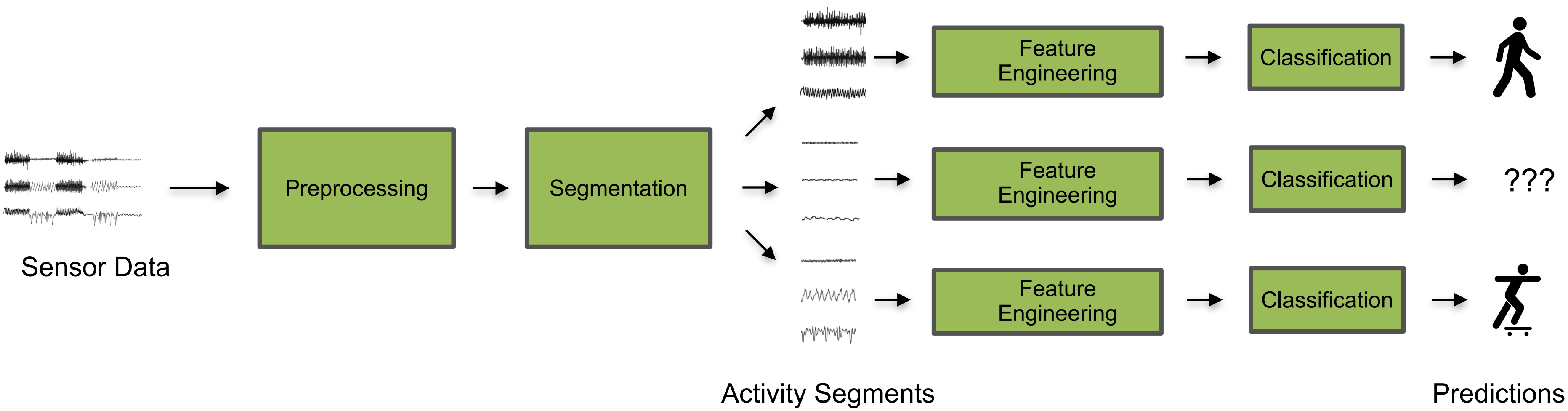
- Changes in shape
- Recurring temporal patterns
- e.g. walking vs. running

TSS for Classification



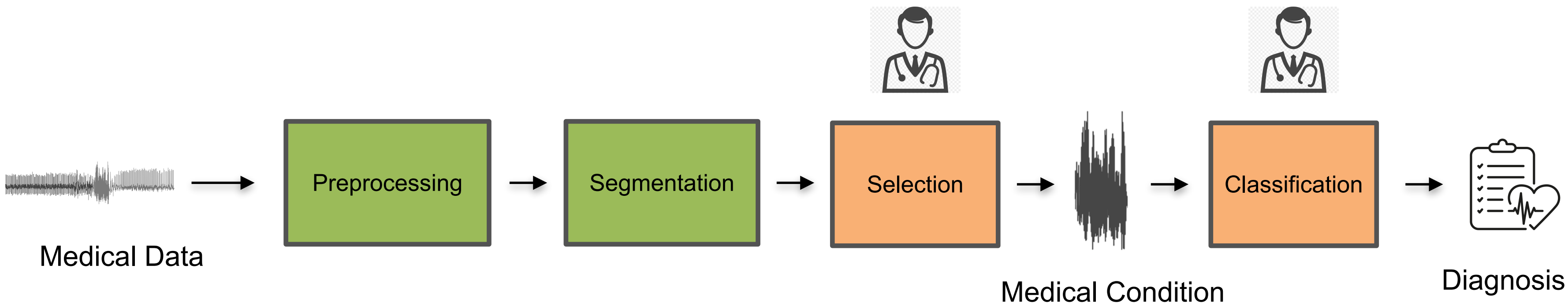
- TSS applied to partition preprocessed TS for classification
- Preprocessing: domain-specific, sensor artefacts, missing values
- Classification of single segments

TSS in Human Activity Recognition (HAR)



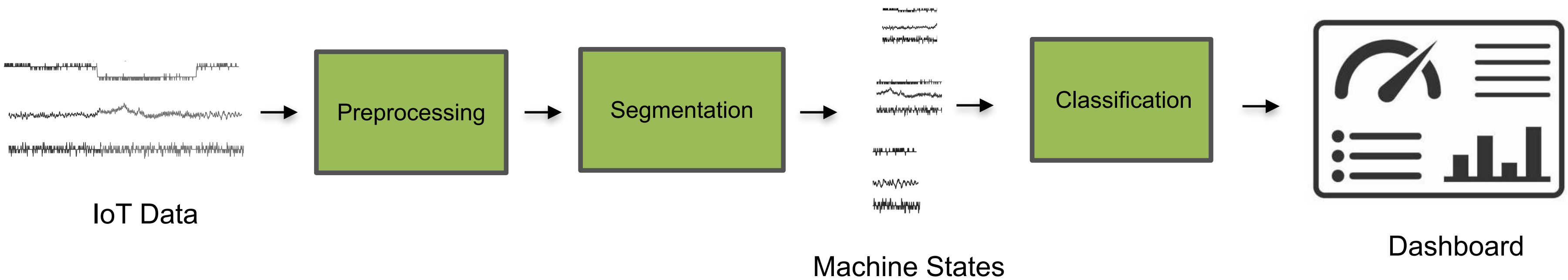
- HAR systems: classify segments of sensor data with activities
- Classification quality depends on TS segment
- TSS in HAR: Partition TS into activity sequence
- Activities typically change in shape

TSS in Medical Condition Monitoring



- Health professionals use medical data to derive medical diagnoses
- e.g. cardiology, sleep evaluation, gait analysis
- TSS divides biomarker measurements into single medical conditions
- Biomarkers typically change in shape

TSS in Smart Manufactoring (IoT)



- Dashboards report process states from machine sensors
- e.g. production lines, power plants, control units
- TSS in IoT: segment sensor data into machine states
- IoT data often change in distribution

Table of Contents

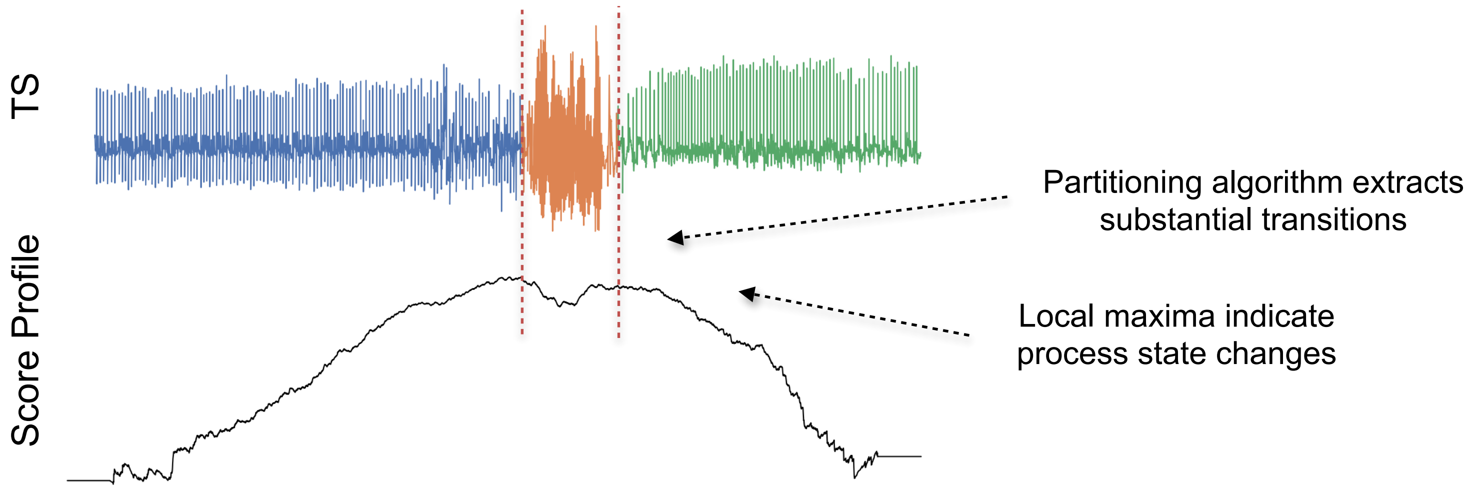
- Time Series Segmentation Task & Use Cases
- **Algorithms**
- Use Cases & Limitations

Algorithmic Advances: Selective Overview

Algorithm	Year	Authors	Publication	Implementation
BOCD	2007	Adams, MacKay	arXiv	
PELT	2012	Killick, Fearnhead, Eckley	Journal of the American Statistical Association	ruptures
AutoPlait	2014	Matsubara, Sakurai, Faloutsos	SIGMOD	
EAgglo	2014	Matteson, James	Journal of the American Statistical Association	aeon
Wild Binseg	2014	Fryzlewicz	The Annals of Statistics	aeon
HOG-1D	2016	Zhao, Itti	WACV	
IGTS	2017	Sadri, Ren	Pervasive and Mobile Computing	aeon
FLOSS	2017	Gharghabi, Ding, Yeh, Kamgar, Ulanova, Keogh	ICDM	aeon
GGS	2019	Hallac, Nystrup, Boyd	Adv. Data Anal. Classif.	aeon
KL-CPD	2019	Chang, Li, Yang, Póczos	ICLR	
ESPRESSO	2020	Deldari, Smith, Sadri, Salim	Interact. Mob. Wearable Ubiquitous Technol.	
Hidalgo	2020	Allegra	Scientific Reports	aeon
TS-CP2	2021	Deldari, Smith, Xue, Salim	WWW	
TIRE	2021	De Ryck, De Vos, Bertrand	IEEE Signal Processing	
ClaSP	2021	Schäfer, Ermshaus, Leser	CIKM	aeon
tGLAD	2023	Imani, Shrivastava,	AALTD	
iCID	2024	Cao, Zhu, Ting, Salim, Li, Yang, Li	JAIR	

Last 15 years have seen a wealth of new TSS algorithms

Common Algorithmic Approach to TSS



- Two main components
    1. Score profile: annotates TS with likelihood of state change
    2. Partitioning algorithm: uses score profile to split TS into segments

Optimisation Problems

Solve: min_{cps} Cost(T, cps) $+$ Penalty(cps)

Search Procedures -> Cost Functions + Constraints

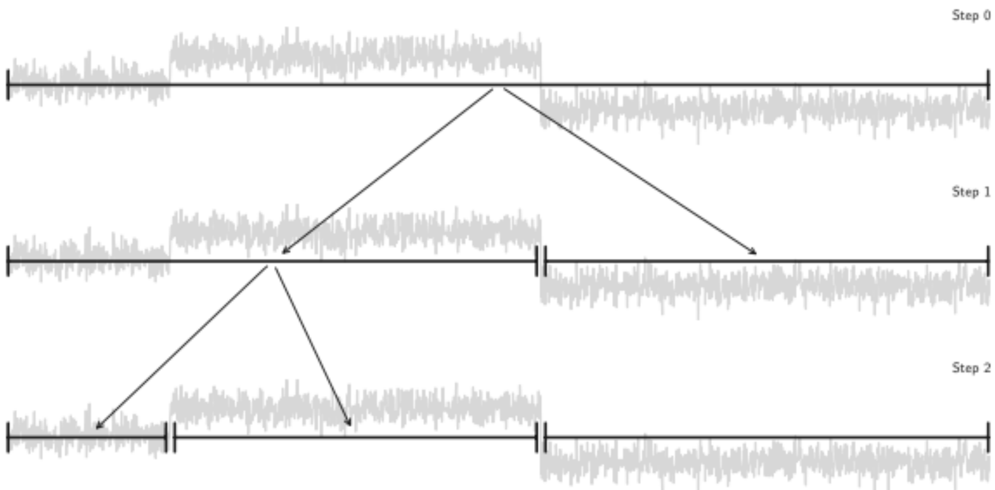
Dyn. Programming
PELT
(Wild-) BinSeg
Window

Maximum Likelihood
Mahalanobis
Rank
Kernel

predefined
AIC
BIC
MDL

- TSS as optimisation problem: minimise summed costs of segments
- Locate potential segments, measure their homogeneity, penalise their amount
- Typical framework for changes in distribution
- Many specific solutions for different signals

Optimisation Problems: Binary Segmentation

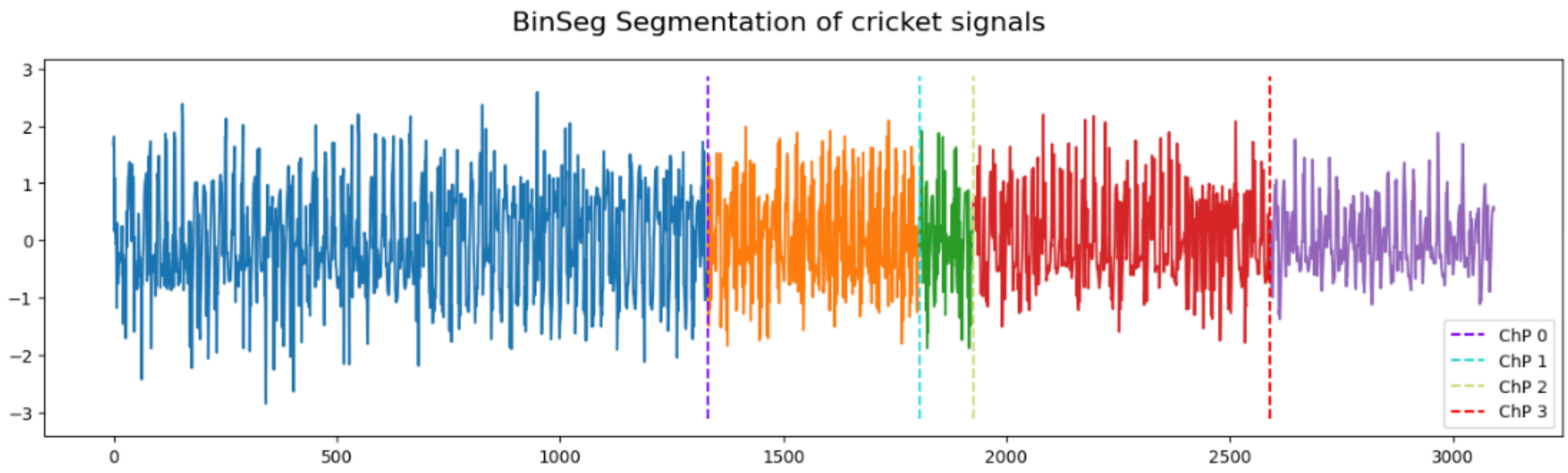


Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. Signal Processing.

- Idea: Recursively split (sub-)signal into two segments
- Split criterion: $argmin_{1<s<|T|}$ Cost$(T_{1,s})$ + Cost$(T_{s+1,|T|})$
- Popular approximation, extensions called circular / wild

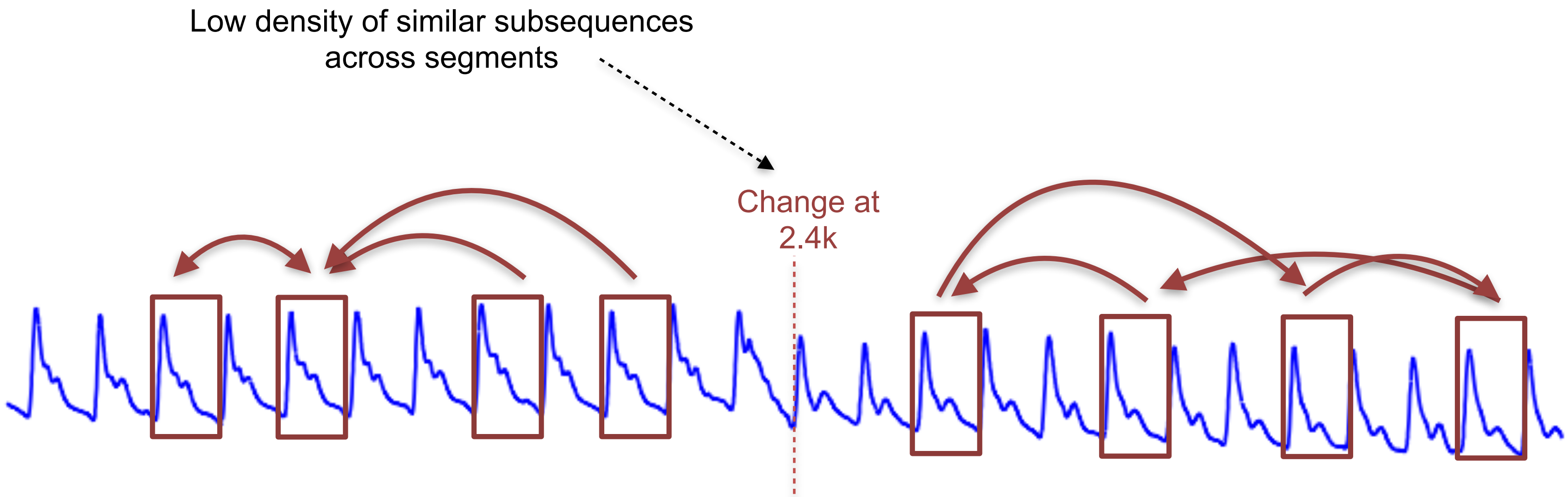
aeon: Binary Segmentation

```
binseg = BinSegmenter(n_cps=len(cps), model="ar")
found_cps = binseg.fit_predict(ts)
plot_series_with_change_points(pd.Series(ts), found_cps, title="BinSeg Segmentation of cricket signals")
```



BinSeg correctly identifies 3 out of 4 CPs

Density of Similar Subsequences

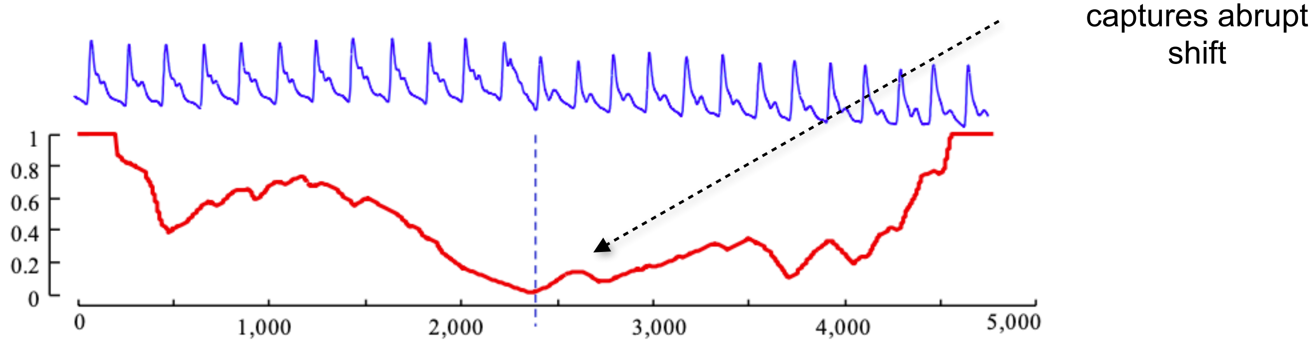


Aterial blood pressure recording
Gharghabi, S., Yeh, C. C. M., Ding, Y., Ding, W., Hibbing, P., LaMunion, S., ... & Keogh, E. (2019).
Domain agnostic online semantic segmentation for multi-dimensional time series. Data mining and knowledge discovery.

- Clustering problem: segments contain mutually similar subsequences
- Change points: Transition between groups of similar subsequences
- Arcs: nearest neighbour relationships (reveal density information)

FLUSS: Arc Curve

TS: Reclining man rotated into standing position



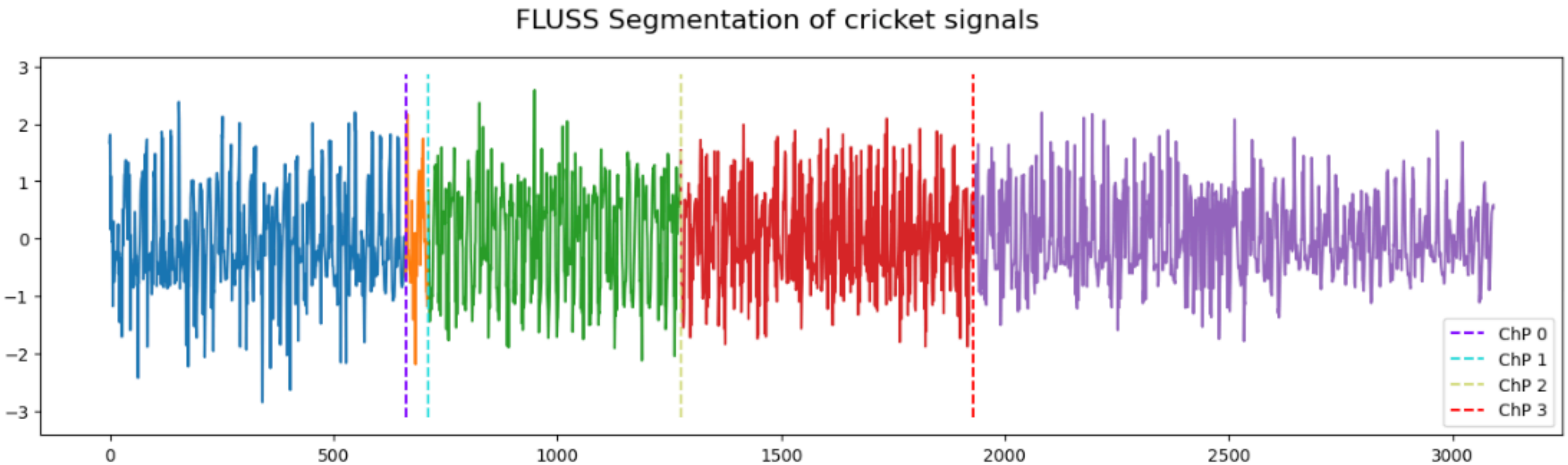
Arc Curve Profile

Gharghabi, S., Yeh, C. C. M., Ding, Y., Ding, W., Hibbing, P., LaMunion, S., ... & Keogh, E. (2019).
Domain agnostic online semantic segmentation for multi-dimensional time series. Data mining and knowledge discovery.

- Arc curve: number of nearest-neighbour crossings
- Idea: Split TS at nearest-neighbour arc curve minima
- Multivariate / streaming version, ESPRESSO extension

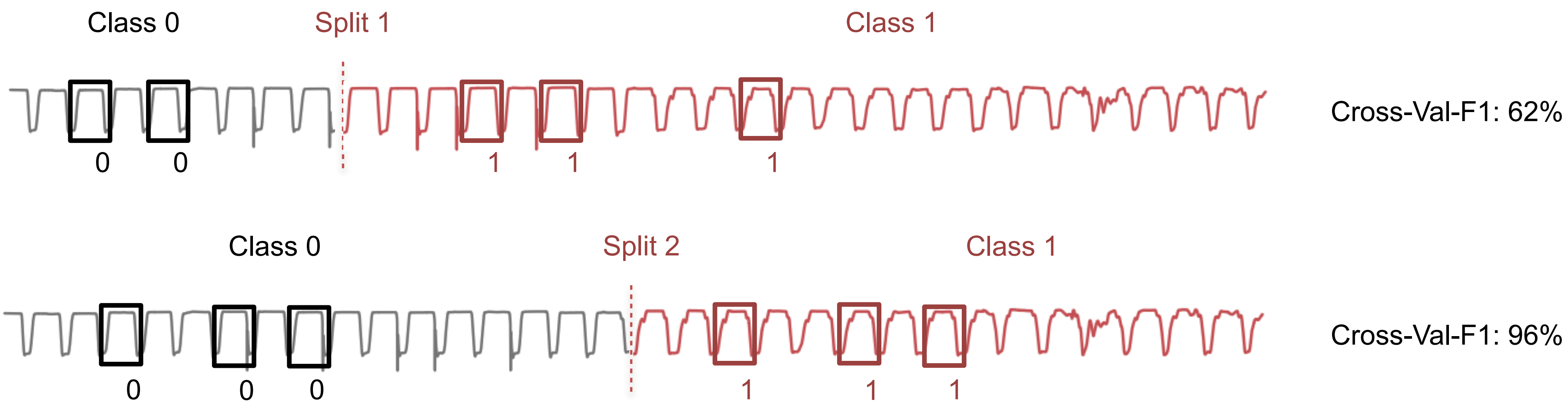
aeon: FLUSS

```
window_size = find_dominant_window_sizes(ts)
fluss = FLUSSSegmenter(period_length=window_size, n_regimes=len(cps)+1)
found_cps = fluss.fit_predict(ts)
plot_series_with_change_points(pd.Series(ts), found_cps, title="FLUSS Segmentation of cricket signals")
```



FLUSS correctly locates 3 out of 4 CPs

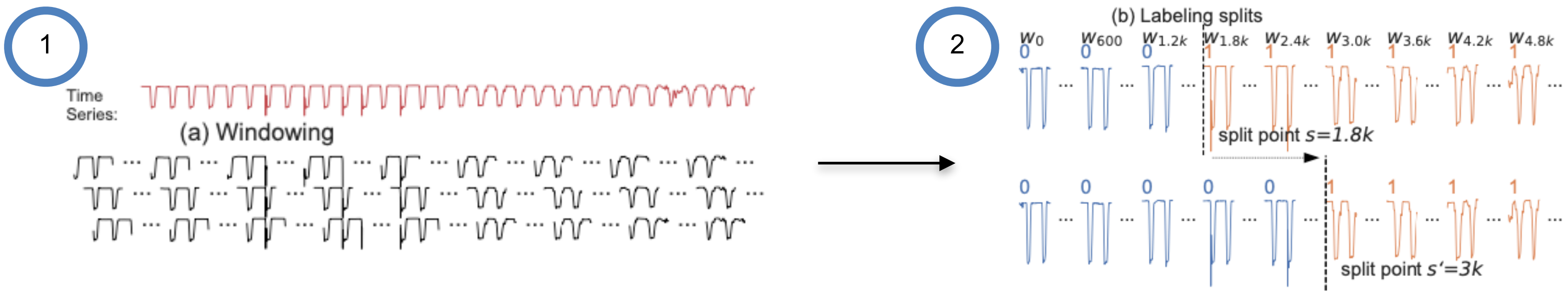
Self-supervised Classification



GunPoint hand motion recording
Ratanamahatana, C. A., & Keogh, E. (2005). Three myths about dynamic time warping data mining.
In Proceedings of the 2005 SIAM international conference on data mining (pp. 506-510). Society for Industrial and Applied Mathematics.

- Self-supervised Classification: TS segments represent different labels
- Find TS splits such that classifier scores high performance on windows
- Leverage development of supervised classification for unsupervised problem

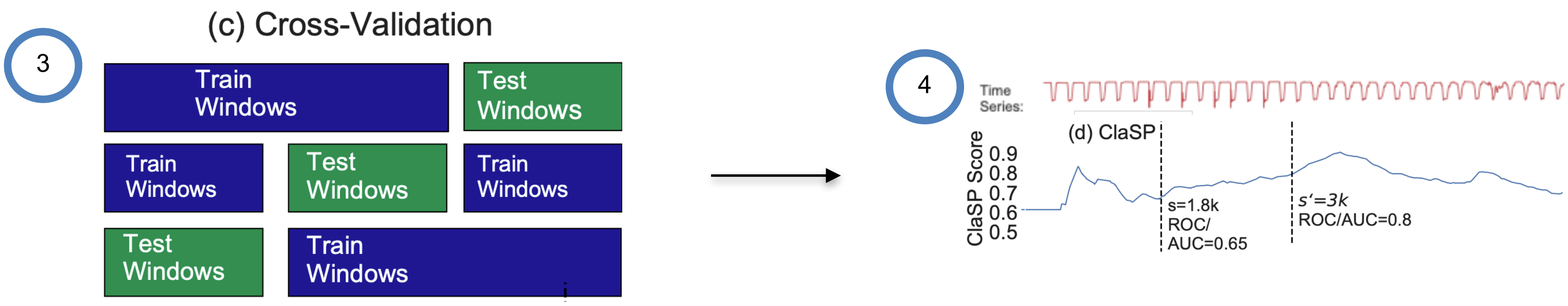
ClaSP: Binary Classification Problem



Schäfer, P., Ermshaus, A., & Leser, U. (2021). Clasp-time series segmentation. In Proceedings of the 30th ACM international conference on information & knowledge management.

- Idea: Create artificial binary subsequence classification problems
- Subsequences either belong to segment with class 0 or 1
- Efficient enumeration of different labelings

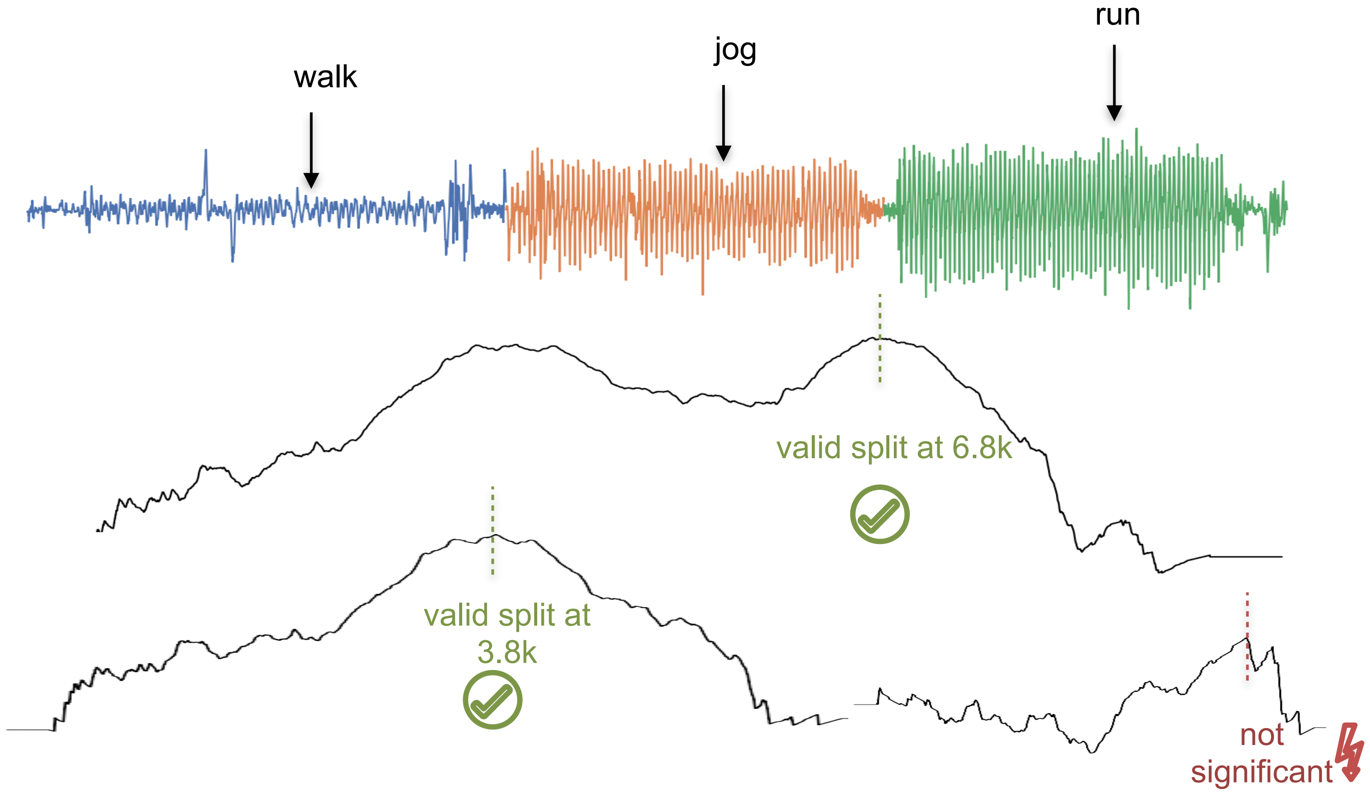
ClaSP: Profile



Schäfer, P., Ermshaus, A., & Leser, U. (2021). Clasp-time series segmentation. In Proceedings of the 30th ACM international conference on information & knowledge management.

- Efficient *k*-NN classifier relabelling, parameter-free
- Profile: cross-validation scores for hypothetical splits
- Streaming / multivariate versions, ensembling extension

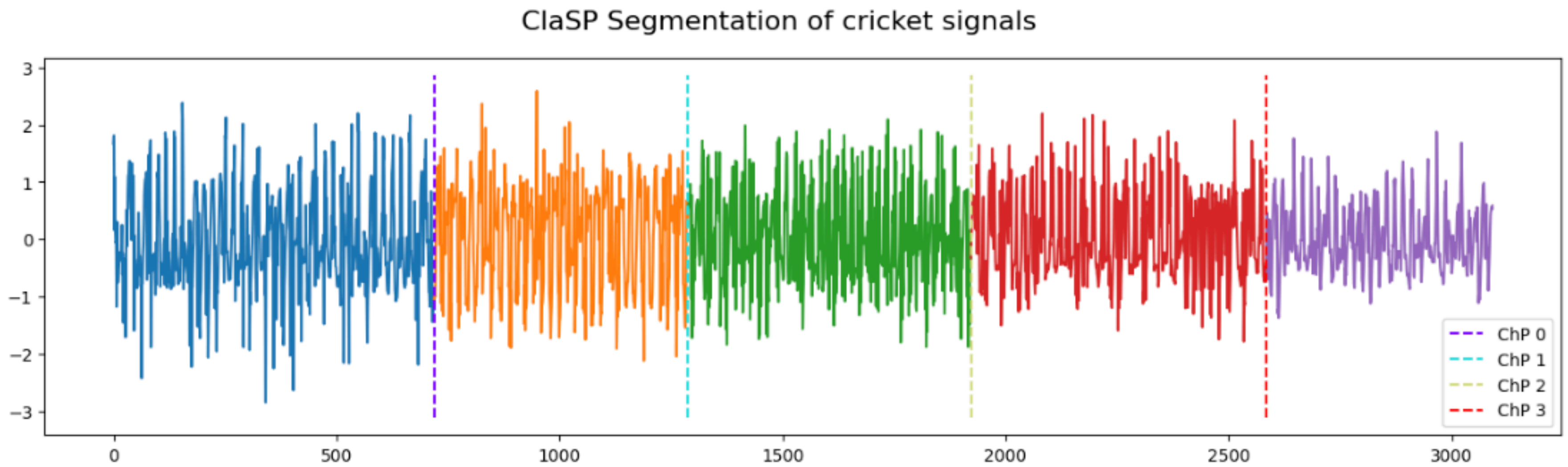
ClaSP: Segmentation



- Local maxima in ClaSP indicate CPs candidates
- Validate if global maximum is CP with hypothesis testing
- Recursively calculate profiles for sub-segments, repeat process

aeon: ClaSP

```
window_size = find_dominant_window_sizes(ts)
clasp = ClaSPSegmenter(period_length=window_size, n_cps=len(cps))
found_cps = clasp.fit_predict(ts)
plot_series_with_change_points(pd.Series(ts), found_cps, title="ClaSP Segmentation of cricket signals")
```

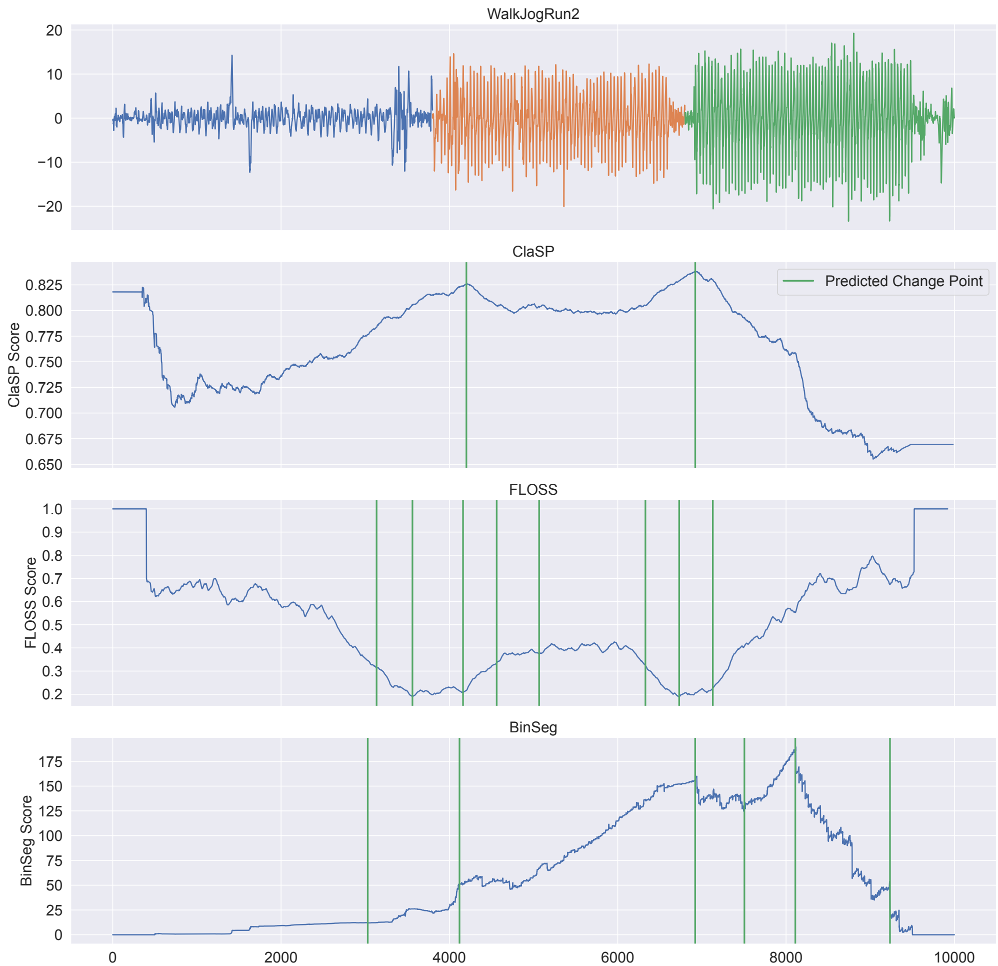


ClaSP correctly identifies all CPs

Table of Contents

- Time Series Segmentation Task & Use Cases
- Algorithms
- **Use Cases & Limitations**

Use Case: Human Activity Segmentation



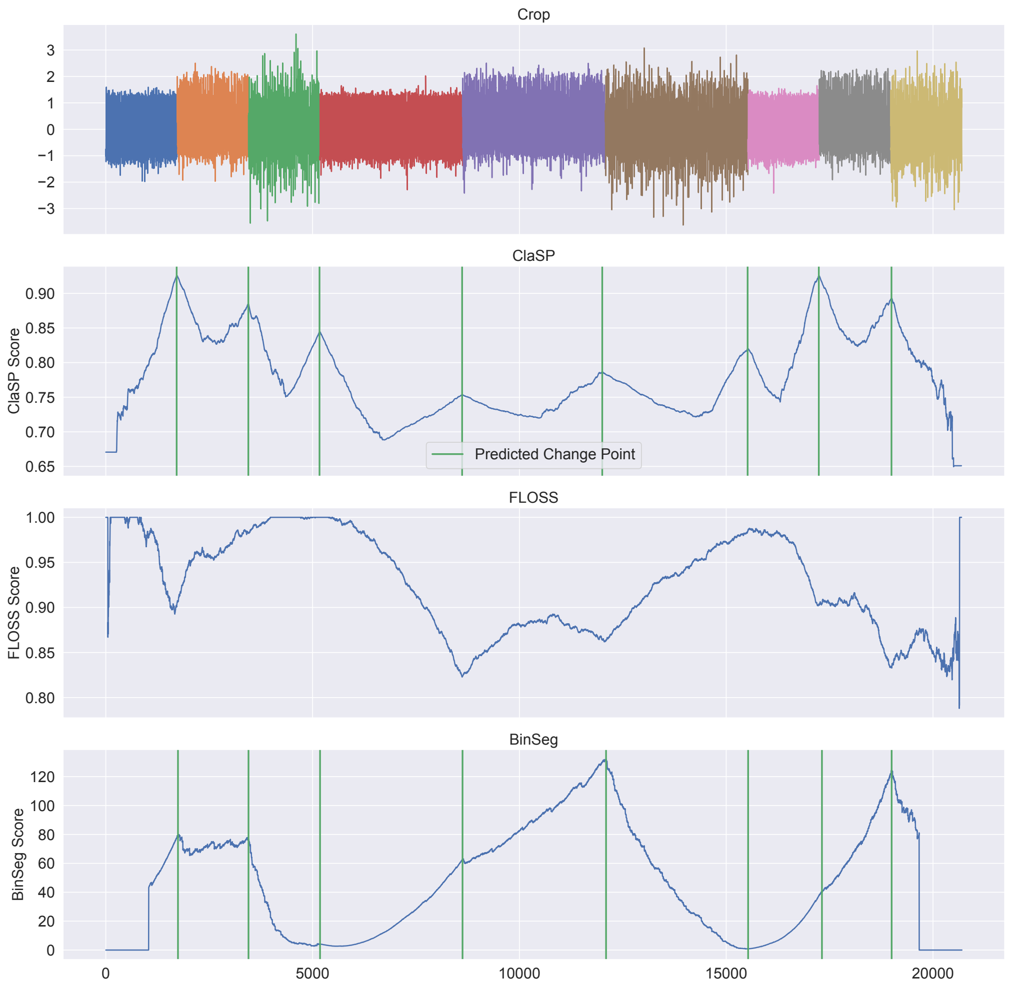
Gyroscope of walking, jogging and running

ClaSP correctly segments TS

FLOSS finds change points and nearby false positives

BinSeg finds change points and noise

Use Case: Satellite Image TS



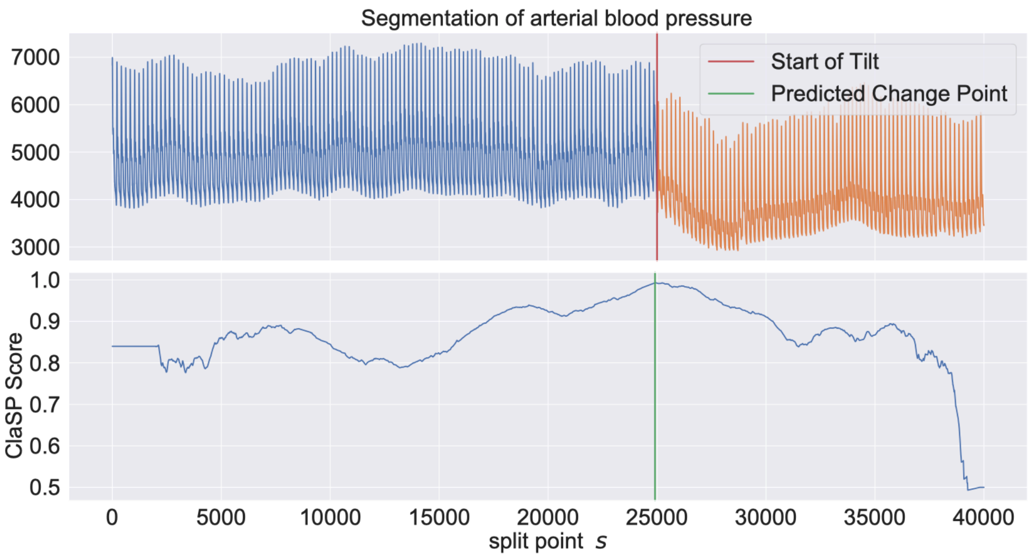
Satellite image TS captures different crops

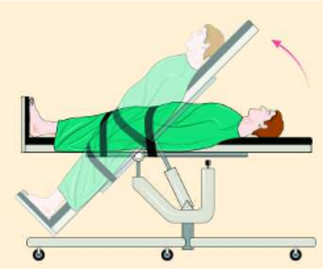
ClaSP correctly segments TS

FLOSS cannot correctly extract CPs from arc curve (threshold of 0.45 not met)

BinSeg correctly segments TS

Open Problems: Gradual Transitions





Gharghabi, S., et al. (2019). Domain agnostic online semantic segmentation for multi-dimensional time series. Data mining and knowledge discovery.

- Arterial blood pressure of volunteer lying on tilt table, which is rapidly turned up
- Sudden rise in blood pressure that slowly drops after reaching upright position
- Current TSS algorithms cannot detect gradual decrease

TSS Extension: State Detection

Ermshaus, A., Schäfer, P., & Leser, U. (2025). CLaP - State Detection from Time Series. *VLDB (accepted)*

- Human activity recording of female performing sport activities
- Problem: Detect latent states (activities) of captured process
- Current algorithms only detect boundaries between segments

Conclusion and Outlook

- TSS: partition TS into homogenous regions
- Advanced preprocessing for classification
- SoTA algorithms: ClaSP, FLOSS, BinSeg
- Challenges: gradual transitions, streaming/multivariate data

TSS is ready to use in aeon

Any questions? Contact me at:
ermshaua@informatik.hu-berlin.de

Thanks for listening!