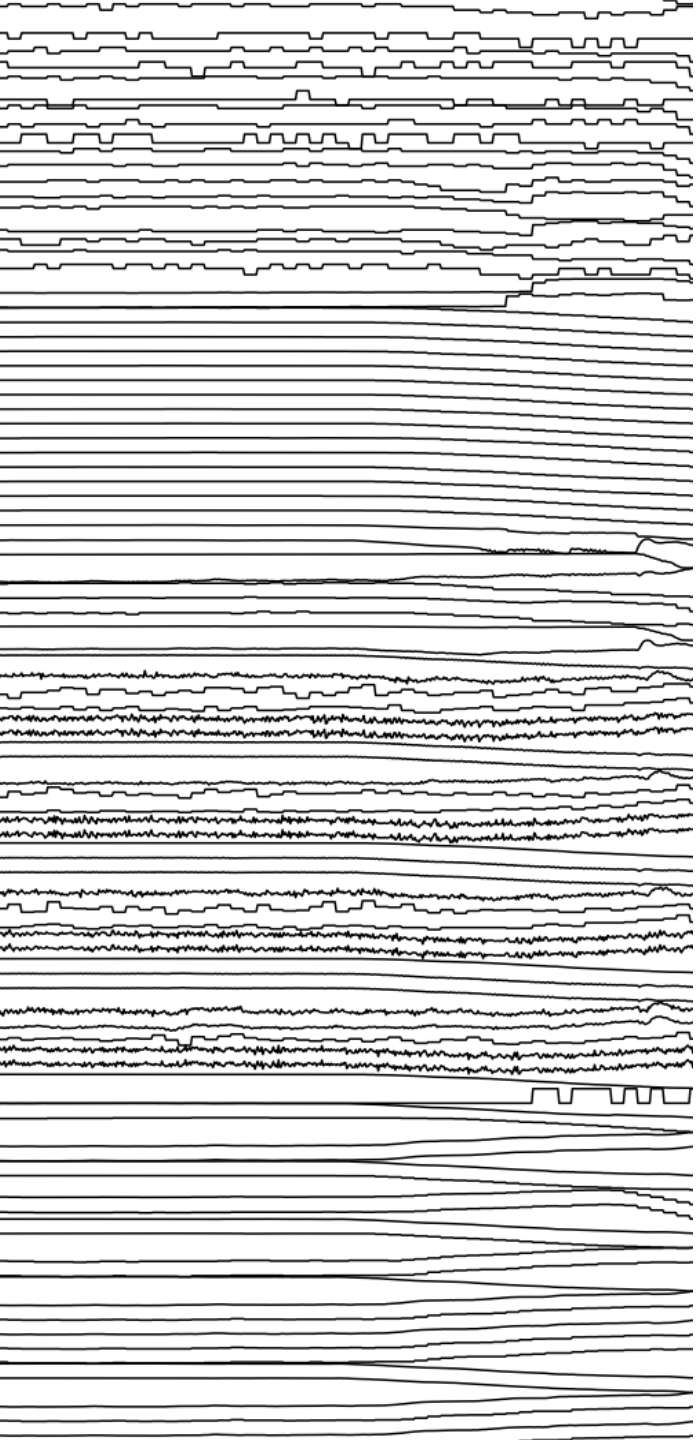
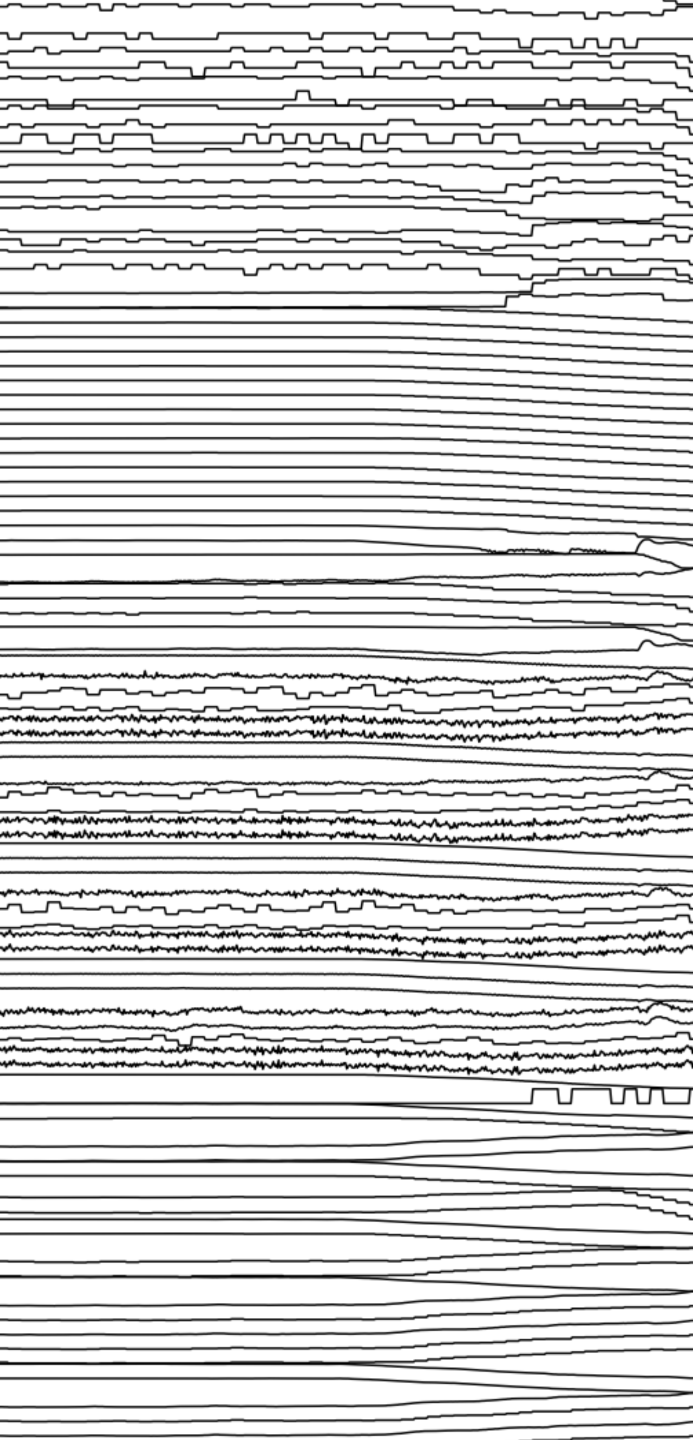


Time Series Anomaly Detection: *Foundations and Practices*

Paul Boniol
Inria, ENS, PSL University
Paul.boniol@inria.fr



Paul Boniol

*I am a researcher at Inria, a member of the VALDA project team, a joint team between **Inria** Paris, **École Normale Supérieure**, and **CNRS**.*

My research interest lies in the intersections between:

- ***Massive time series** analytics and management systems.*
- *Unsupervised and supervised **anomaly detection** methods for large time series.*
- ***Machine learning** for time series analytics.*

paul.boniol@inria.fr

https://boniolp.github.io/

Paul Boniol

ACM SIGMOD Blog

Paul Boniol - Themis Palpanas

JULY 16, 2024

TIME SERIES ANOMALY DETECTION

Time Series

What it is, how it works, where we are, and where we are heading

Anomaly detection is an important problem in data analytics with applications in many domains. In recent years, there has been an increasing interest in anomaly detection tasks applied to time series. In this post, we take a holistic view on anomaly detection in time series, starting from the core definitions and taxonomies related to time series and anomaly types, to a brief description of the anomaly detection methods proposed by different communities in the literature. We conclude with the recent benchmark and evaluation effort proposed recently in the data management community, and new perspectives and research direction.

Inria ENS ÉCOLE NORMALE SUPÉRIEURE | PSL

Introduction: *Time series are Everywhere*

Energy Production

Edf.fr: tinyurl.com/yc7x5xje

Astrophysics

Virgo: https://www.virgo-gw.eu/

Medicine

tinyurl.com/39dx2us4

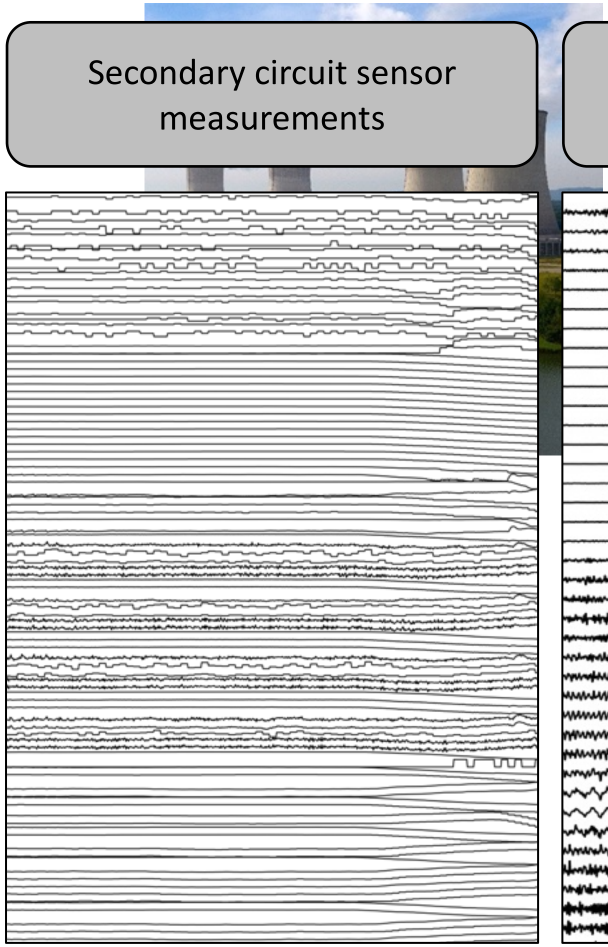
Volcanology

tinyurl.com/ybcttmfz

Introduction: *Time series are Everywhere*

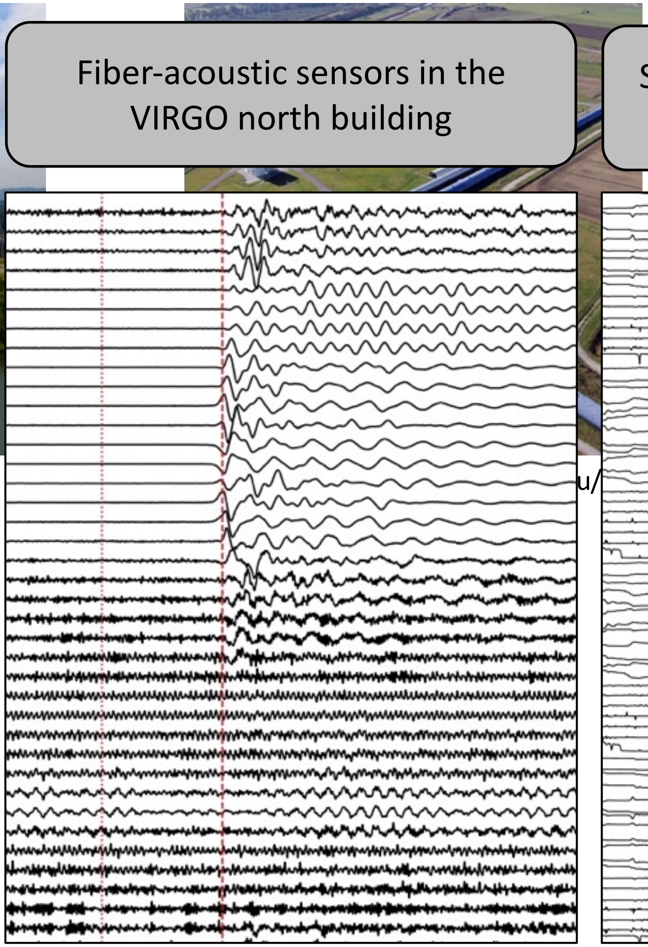
Energy Production

Secondary circuit sensor measurements



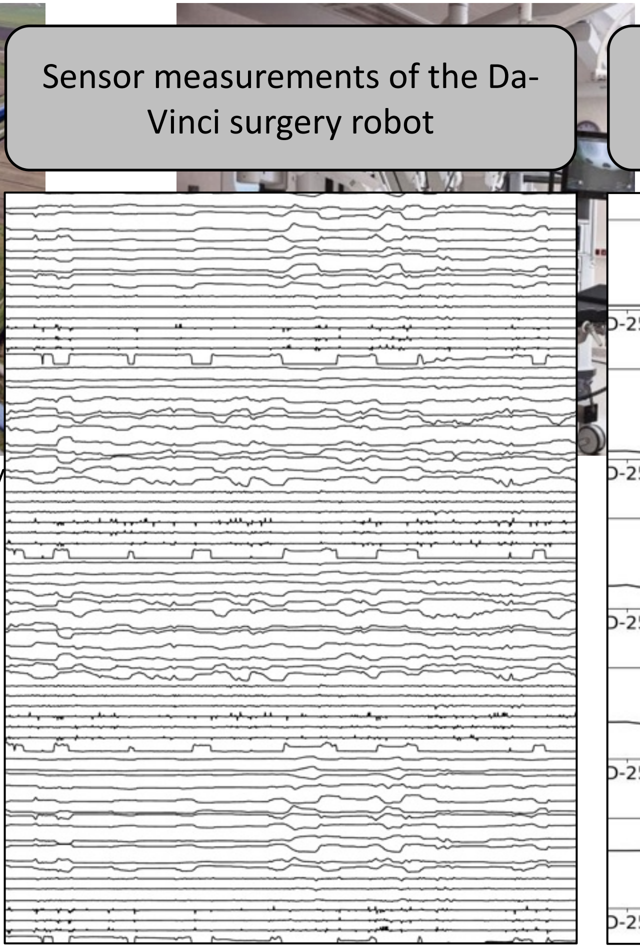
Astrophysics

Fiber-acoustic sensors in the VIRGO north building



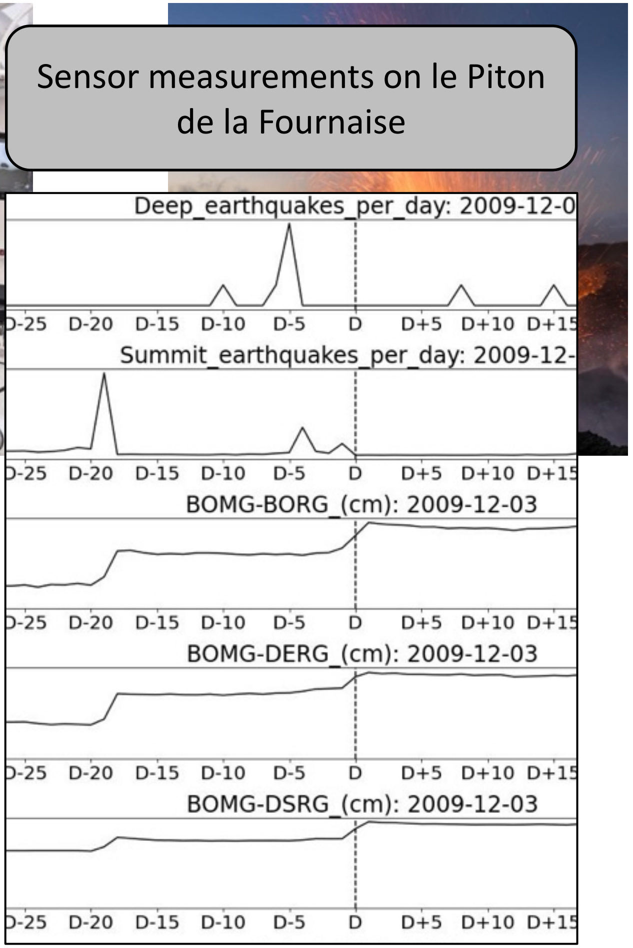
Medicine

Sensor measurements of the Da-Vinci surgery robot



Volcanology

Sensor measurements on le Piton de la Fournaise

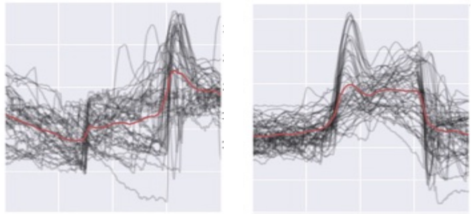


Introduction: *with Important Challenges*

Energy Production

Secondary circuit sensor measurements

Identification of precursors of feed-water pumps vibrations



Astrophysics

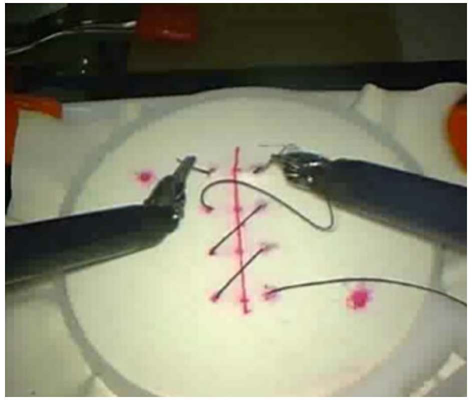
Fiber-acoustic sensors in the VIRGO north building

Noise detection in VIRGO interferometer north building

Medicine

Sensor measurements of the Da-Vinci surgery robot

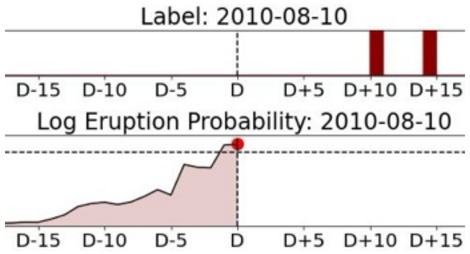
Unusual surgeons gestures detection



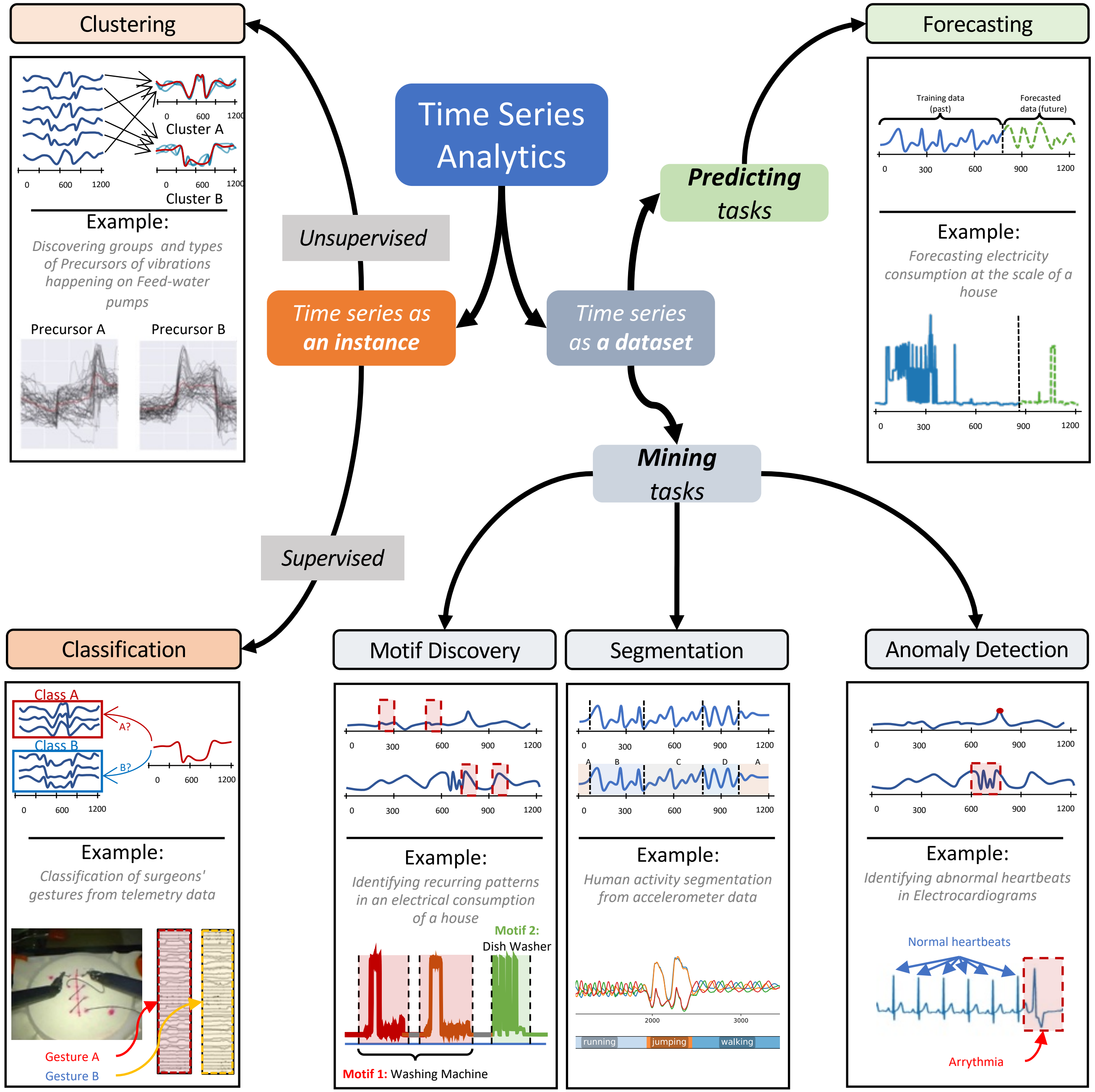
Volcanology

Sensor measurements on le Piton de la Fournaise

Detection of abnormal events on the volcano for predicting eruptions

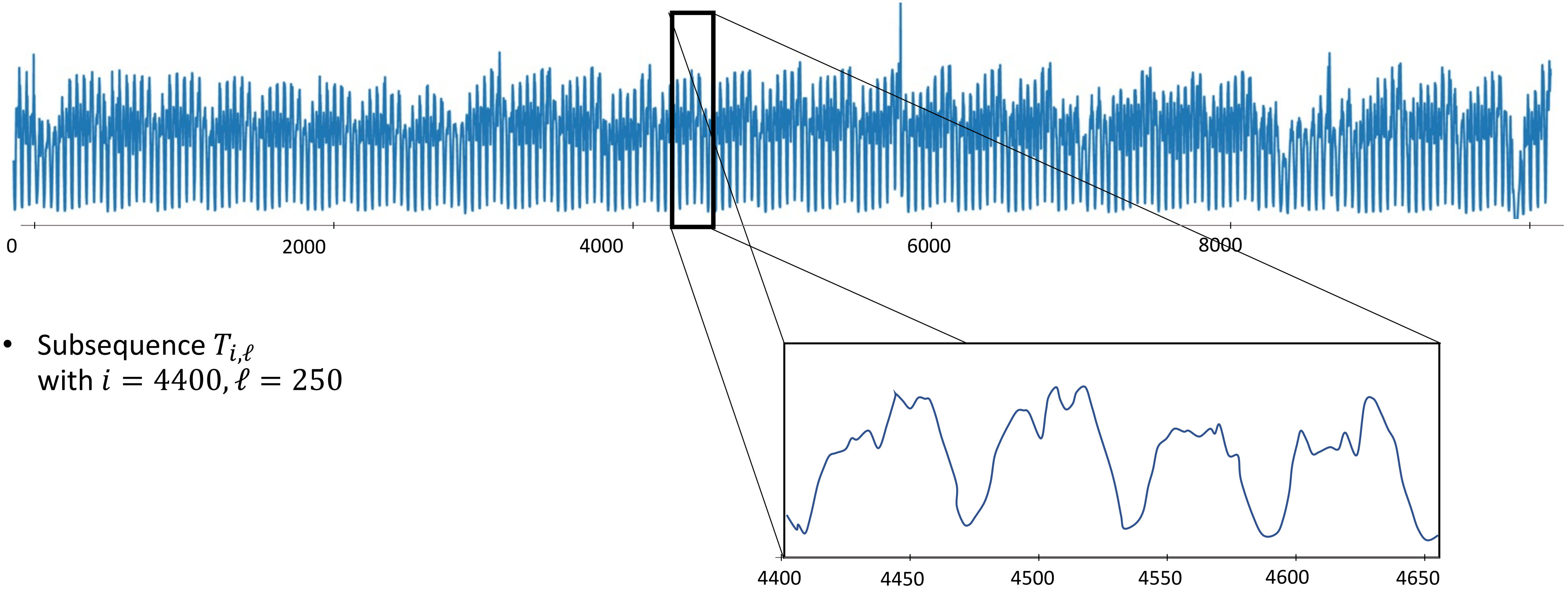


Introduction:
Time Series Analytics



Introduction: *Anomaly Detection in Time Series*

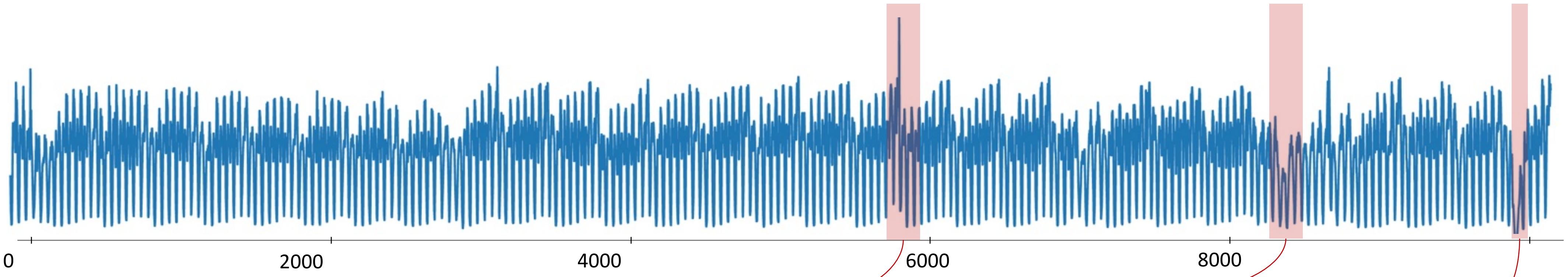
- Time series T *(example : number of taxi passengers in New York City)*



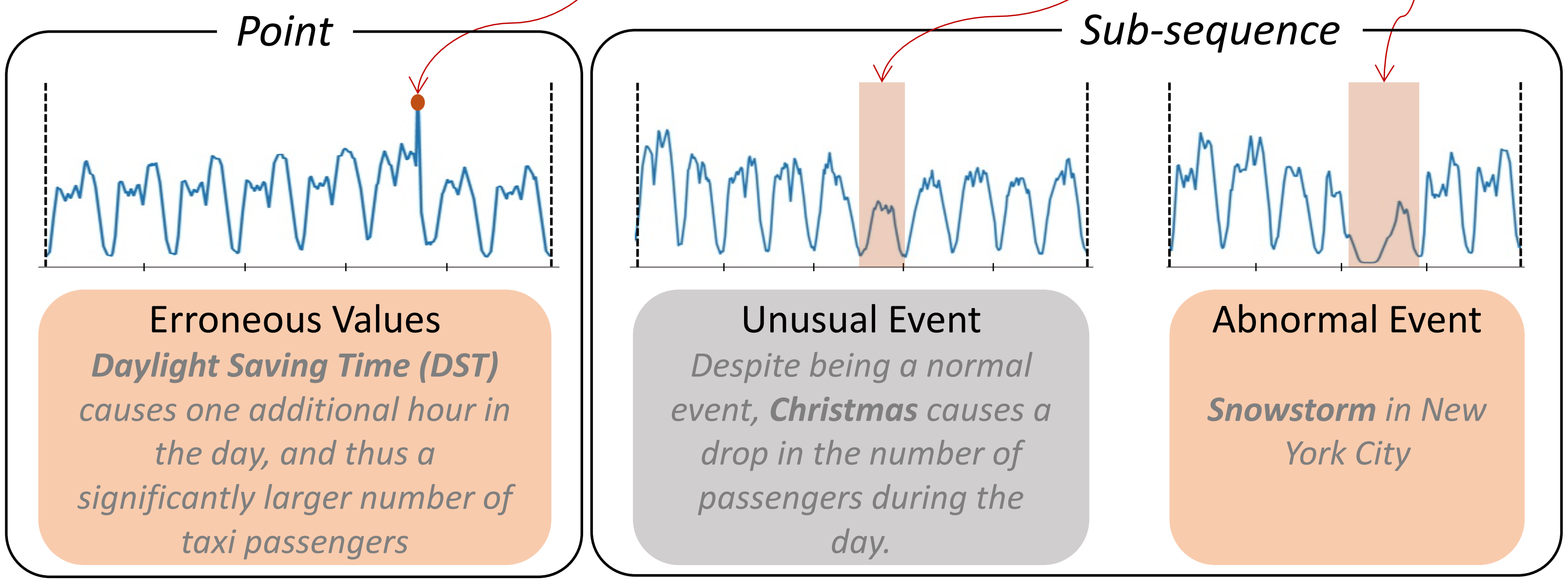
- Subsequence $T_{i,\ell}$
  with $i = 4400, \ell = 250$

Introduction: *Anomaly Detection in Time Series*

- Time series T *(example : number of taxi passengers in New York City)*



- *Anomaly: rare point or sequence (of a given length) potentially non-desired*



Introduction: *Anomaly Detection in Time Series*

- Time series T *(example : number of taxi passengers in New York City)*

Part I

General Overview of Time Series Anomaly Detection

What it is

Part II

Time Series Anomaly Detection Methods

How it works

Part III

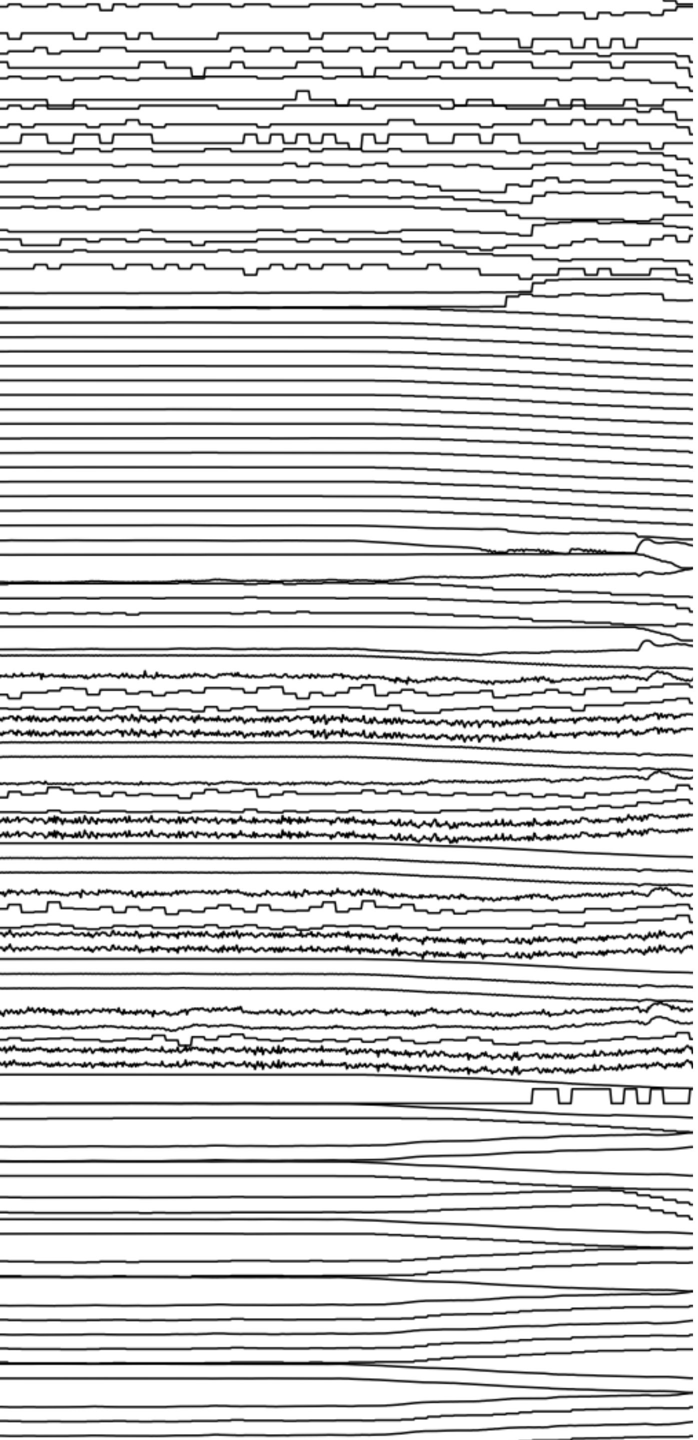
Benchmark, Evaluation Measures, and Automatic Solutions

Where we are

Erroneous Values
Daylight Saving Time (DST) *causes one additional hour in the day, and thus a significantly larger number of taxi passengers*

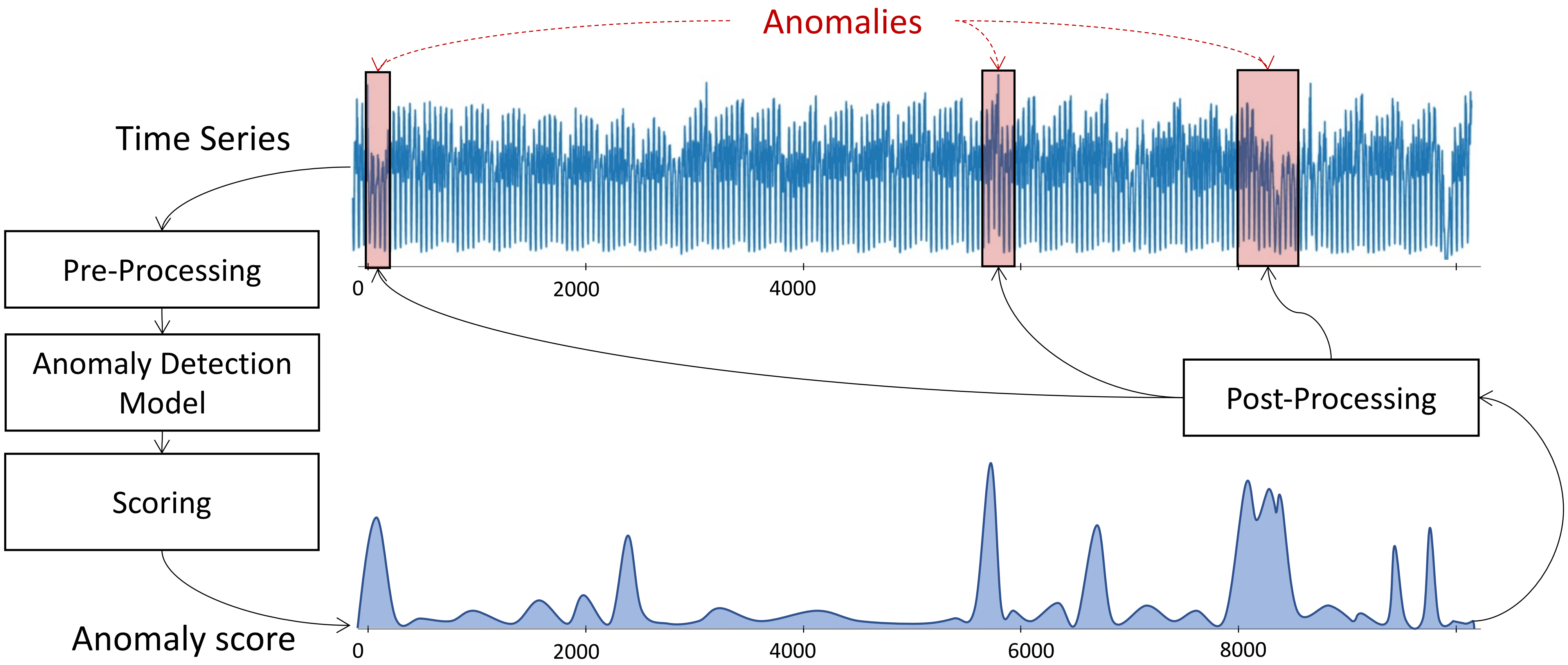
Unusual Event
Despite being a normal event, ***Christmas*** *causes a drop in the number of passengers during the day.*

Abnormal Event
Snowstorm *in New York City*



Part I.
Time Series Anomaly Detection

Anomaly Detection methods: *A taxonomy*



Anomaly D

Time Series

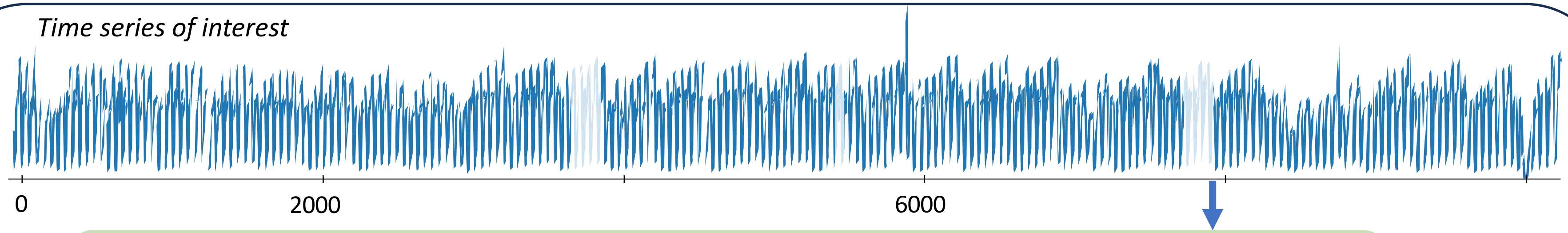
Pre-Processing

Anomaly Detection Model

Scoring

Anomaly score

Time series of interest



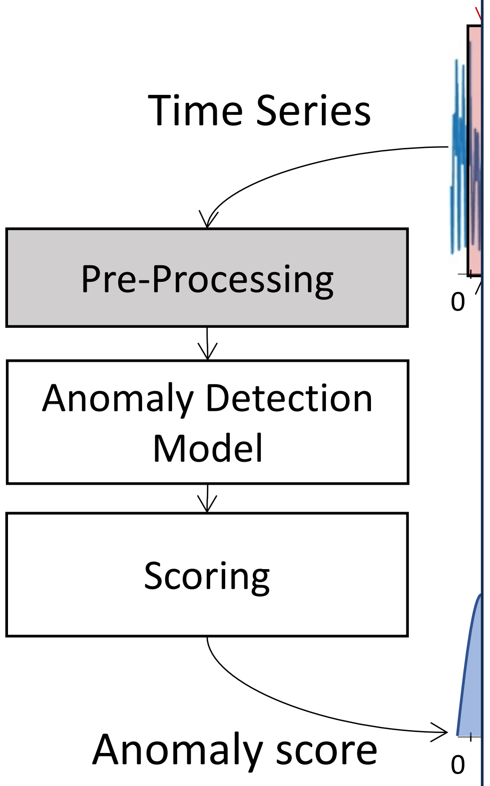
Normalization: *Z-normalization* | *Min-max normalization* | *Denoising* | *Detrending* | ...

Missing Values imputation: *Simple Imputation* (*Mean fill, Forward/Backward fill*) | *Interpolation* (*Linear, Polynomial, etc*) | ...

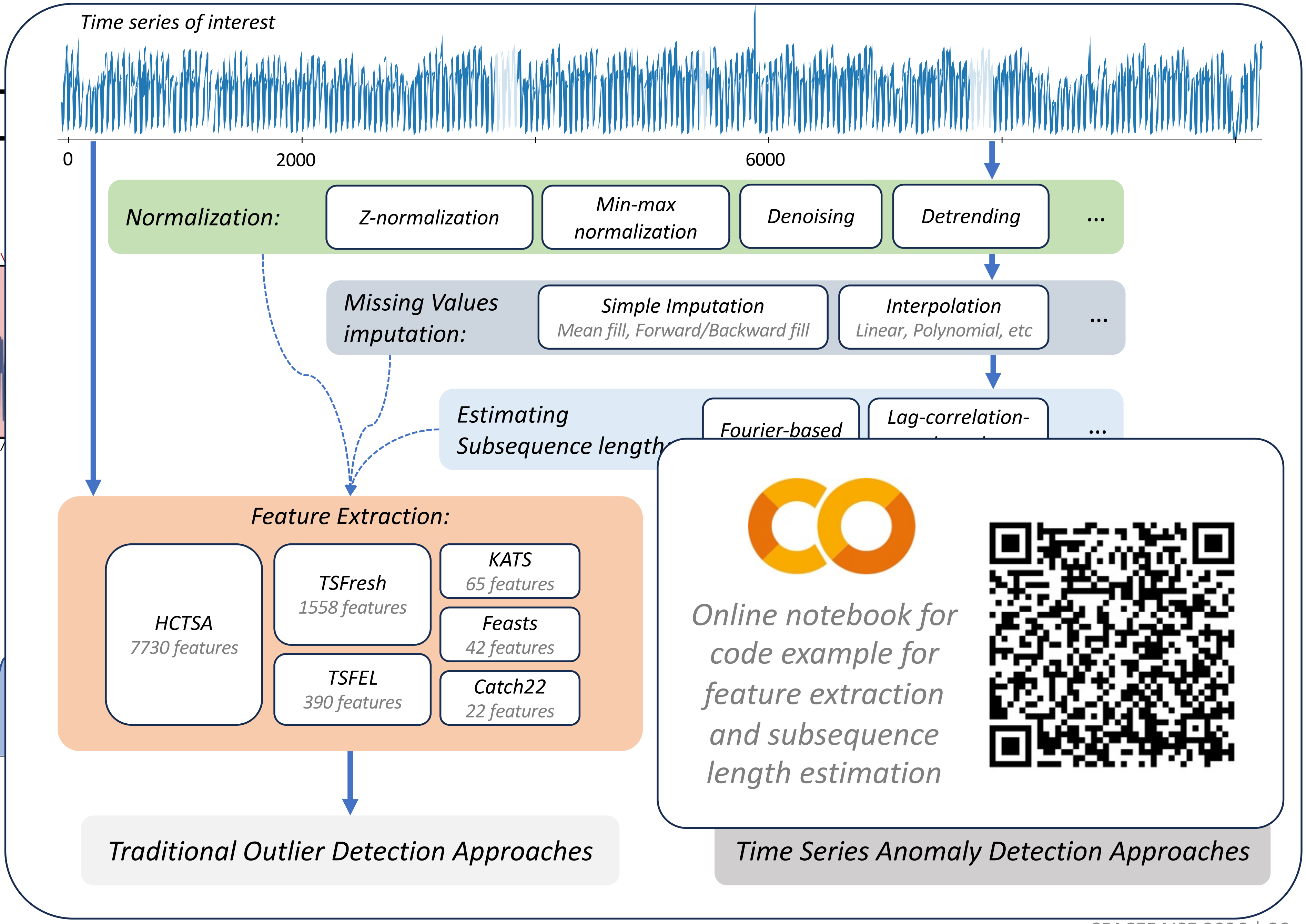
Estimating Subsequence length: *Fourier-based* | *Lag-correlation-based* | ...

Time Series Anomaly Detection Approaches

Anomaly D



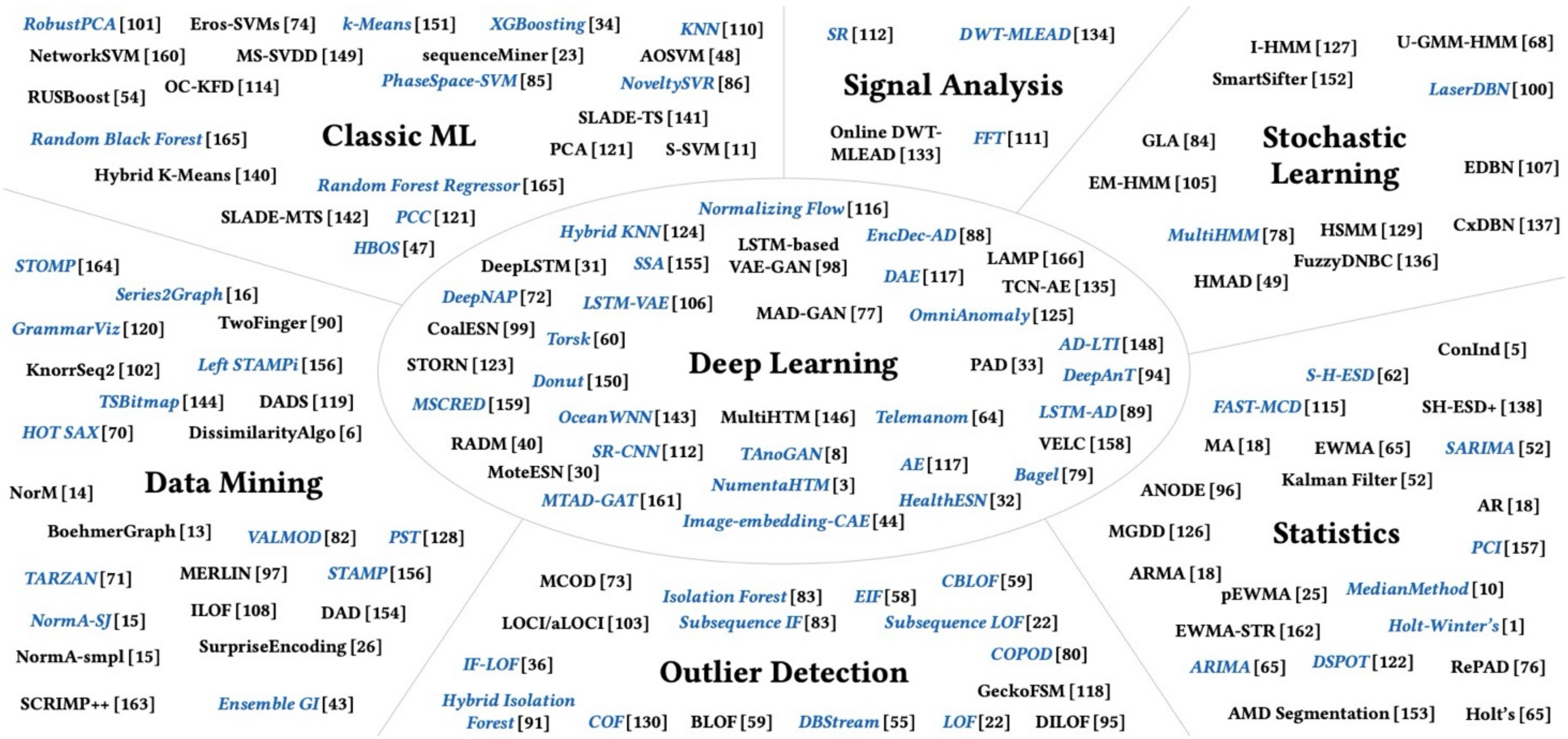
Time Series
Pre-Processing
Anomaly Detection Model
Scoring
Anomaly score
0
0



Time series of interest
0
2000
6000
Normalization:
Z-normalization
Min-max normalization
Denoising
Detrending
...
Missing Values imputation:
Simple Imputation
Mean fill, Forward/Backward fill
Interpolation
Linear, Polynomial, etc
...
Estimating Subsequence length
Fourier-based
Lag-correlation-
...
Feature Extraction:
HCTSA
7730 features
TSFresh
1558 features
TSFEL
390 features
KATS
65 features
Feasts
42 features
Catch22
22 features
Traditional Outlier Detection Approaches
Time Series Anomaly Detection Approaches
Online notebook for code example for feature extraction and subsequence length estimation

Anomaly Detection methods: *A taxonomy*

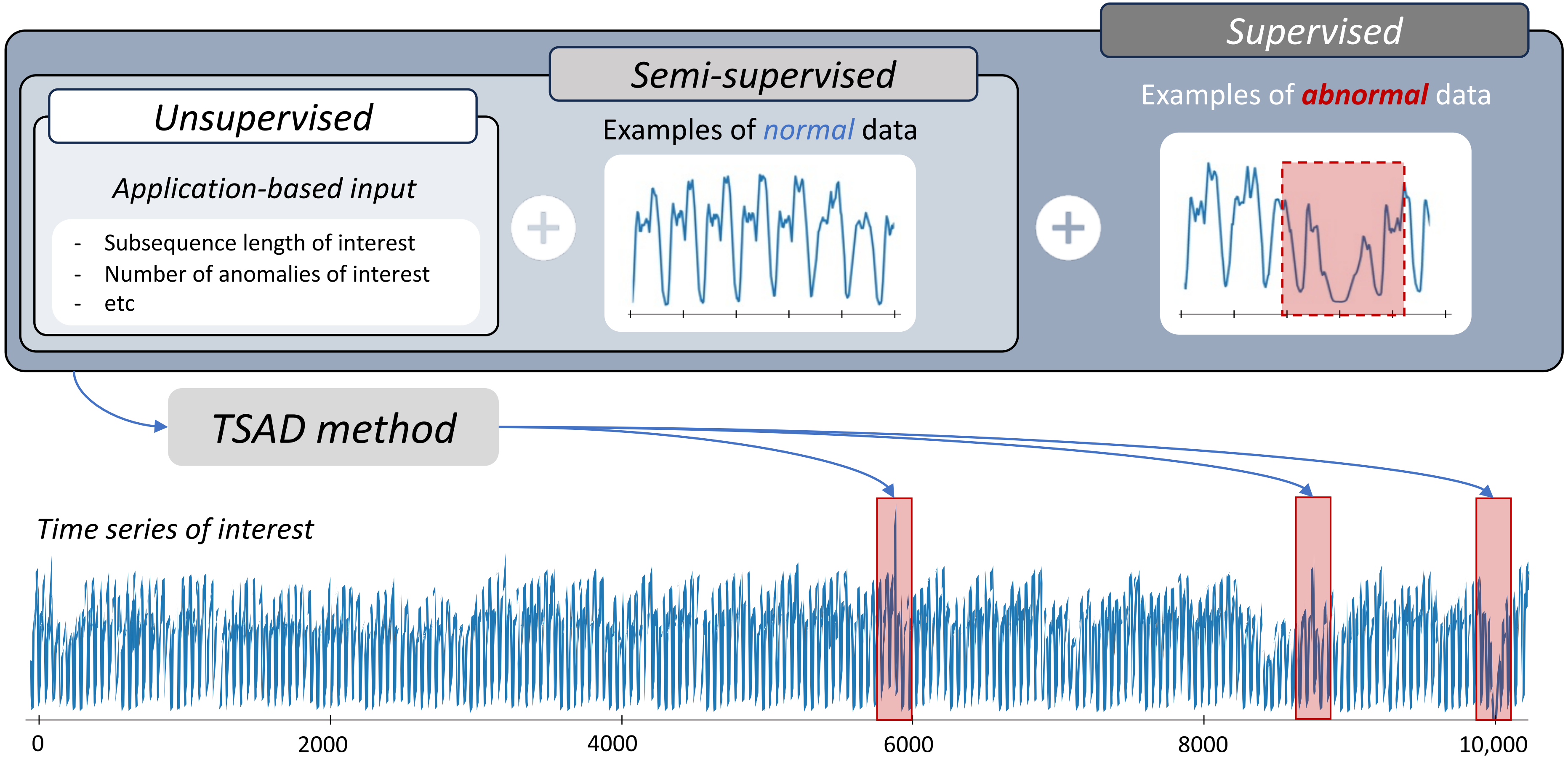
By domains [5] …



[5] Sebastian Schmidl, Phillip Wenig, and Thorsten Papenbrock. 2022. Anomaly detection in time series: a comprehensive evaluation. Proc. VLDB Endow. 15, 9 (May 2022), 1779–1797.

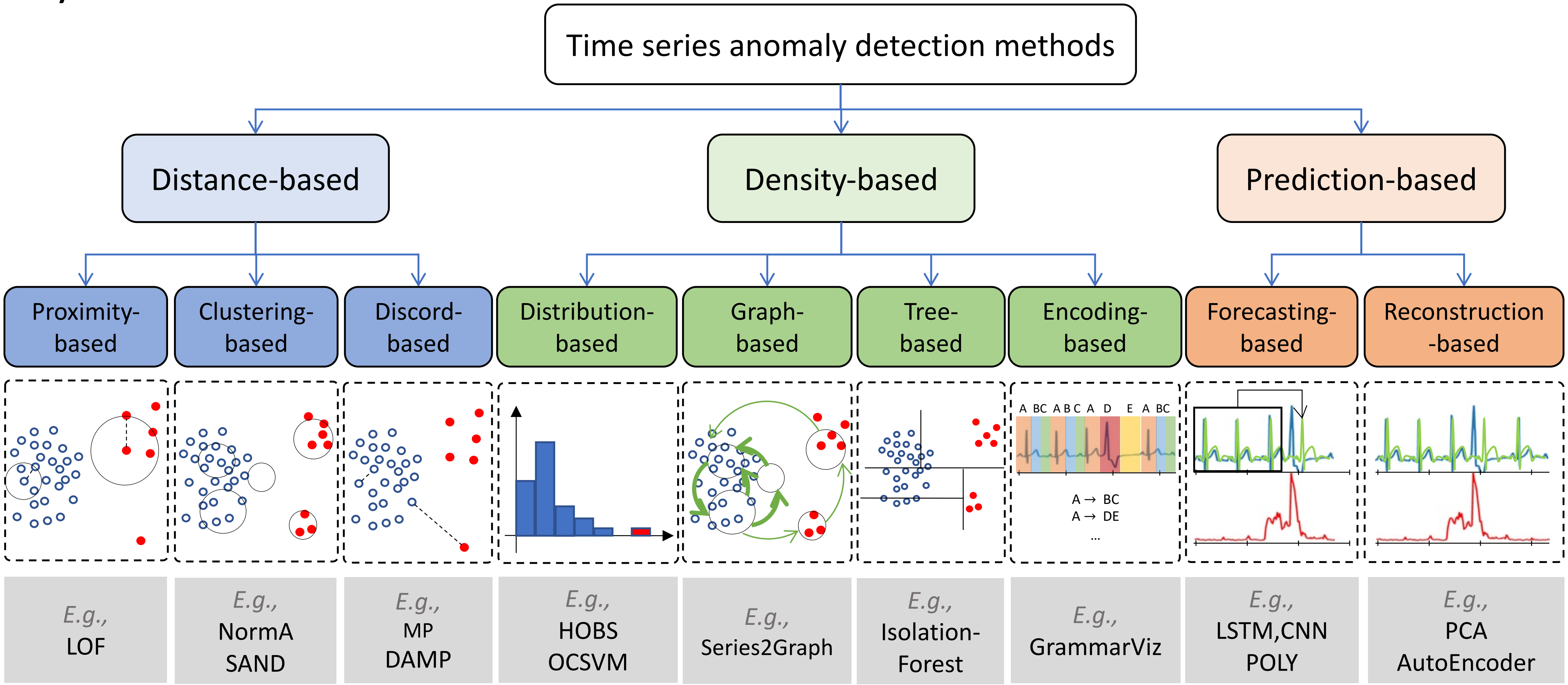
Anomaly Detection methods: *A taxonomy*

By inputs…

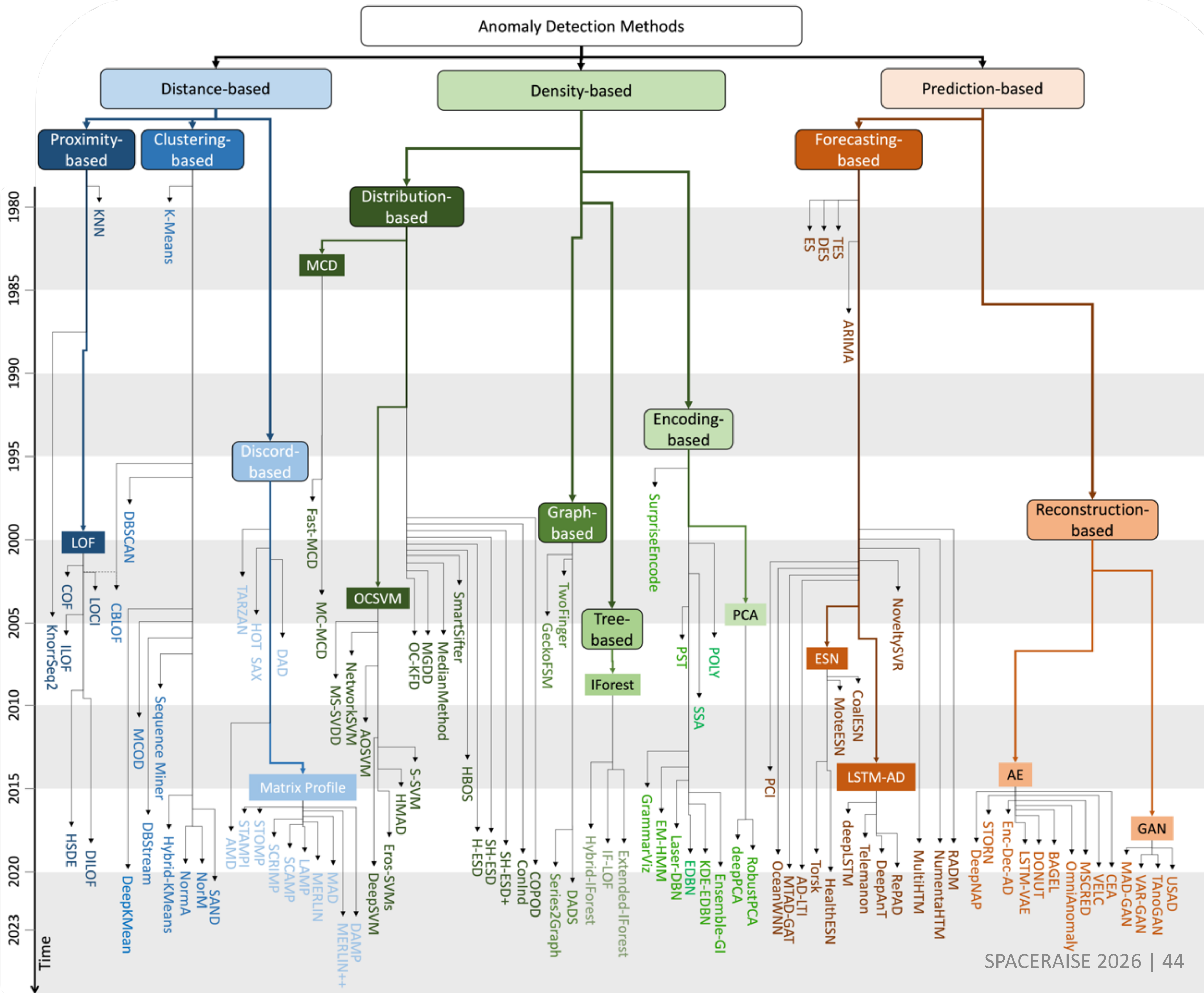


Anomaly Detection methods: *A taxonomy*

By methods…



Anomaly Detection methods:
A taxonomy

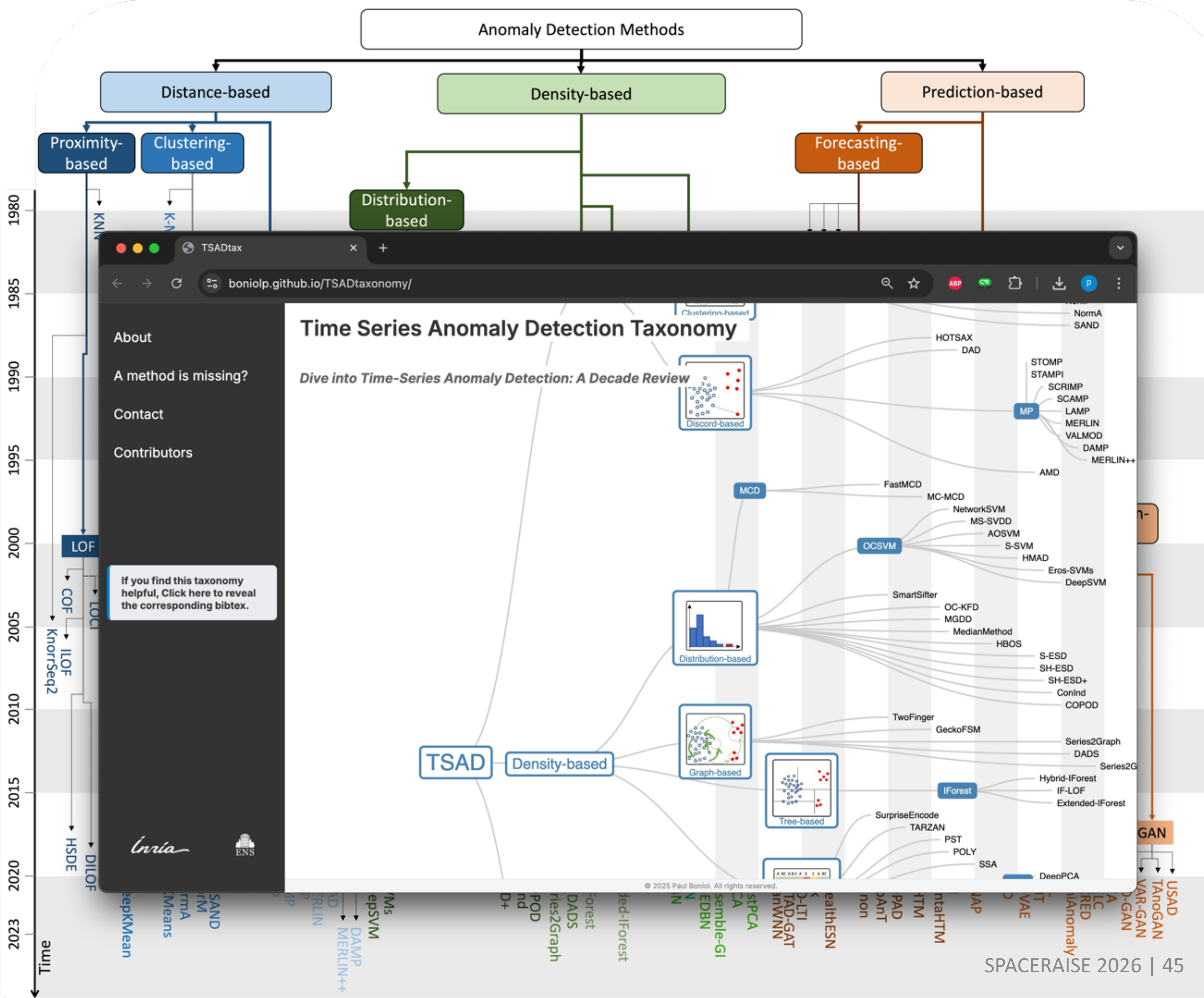


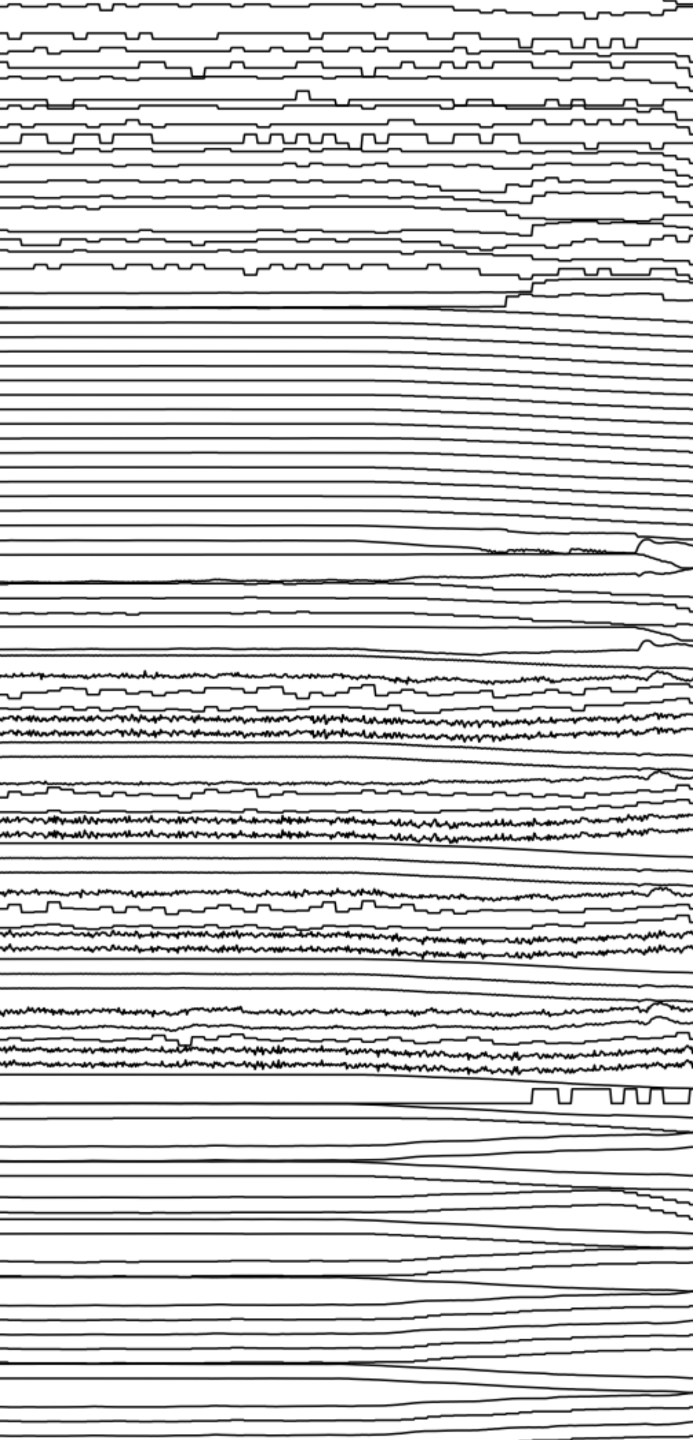
Anomaly Detection methods:
A taxonomy

Online interactive Taxonomy

https://boniolp.github.io/TSADtaxonomy/

https://github.com/boniolp/TSADtaxonomy

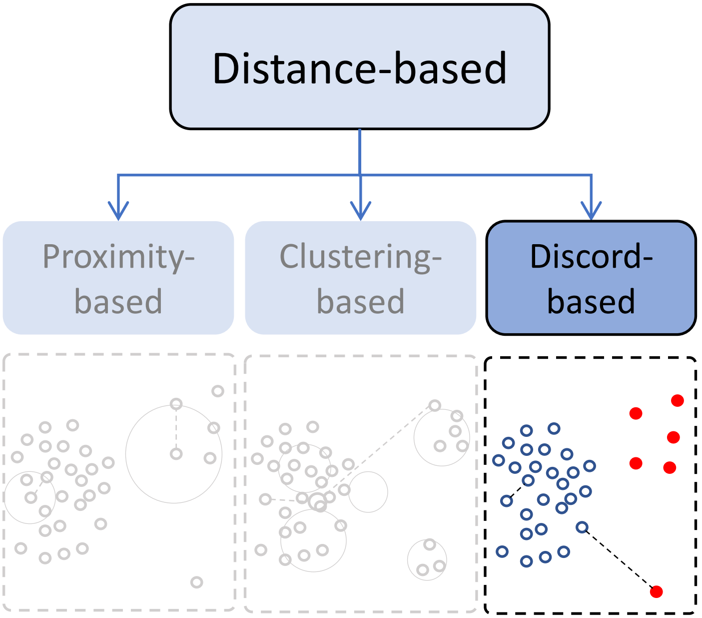


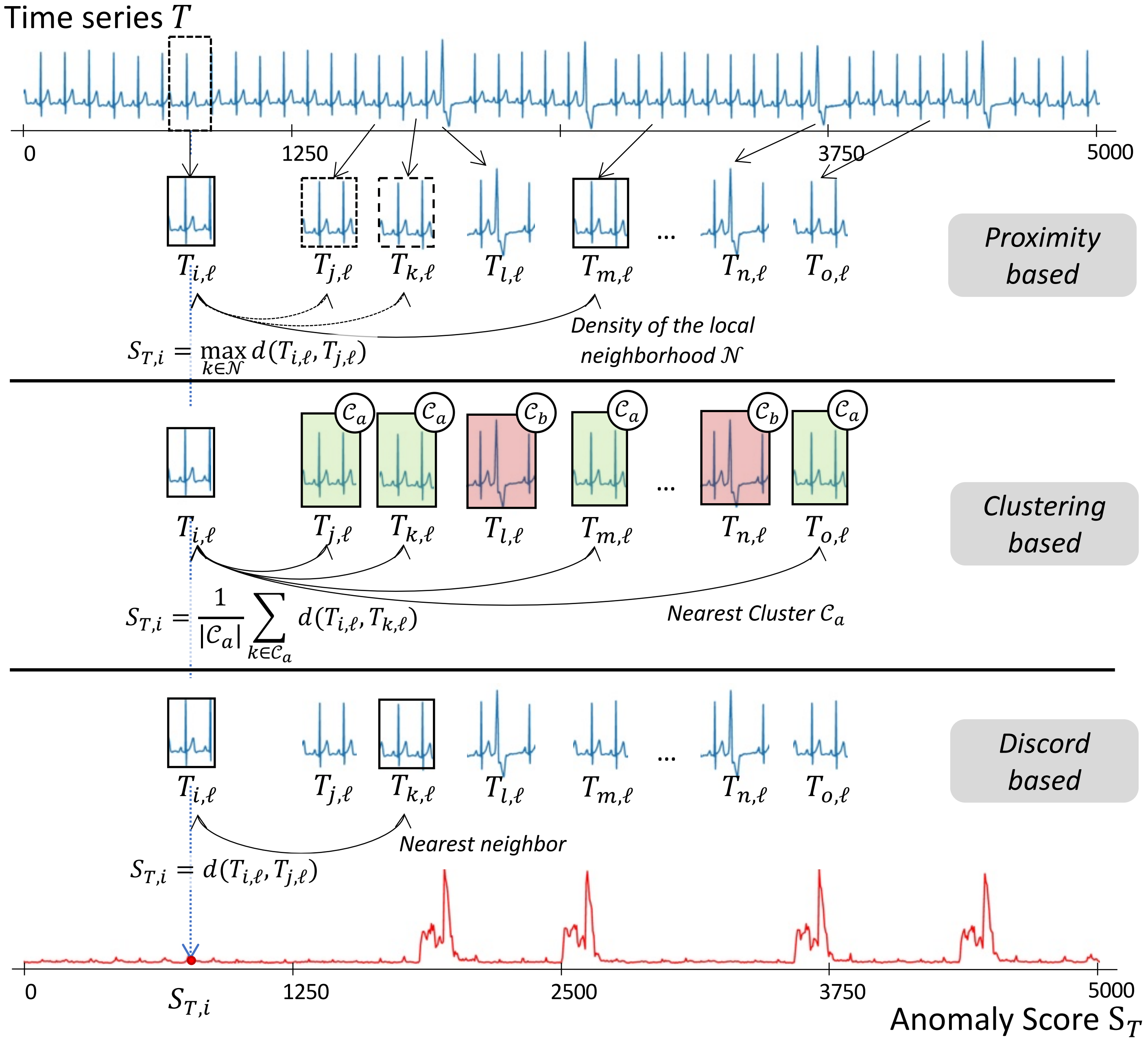


Part II.

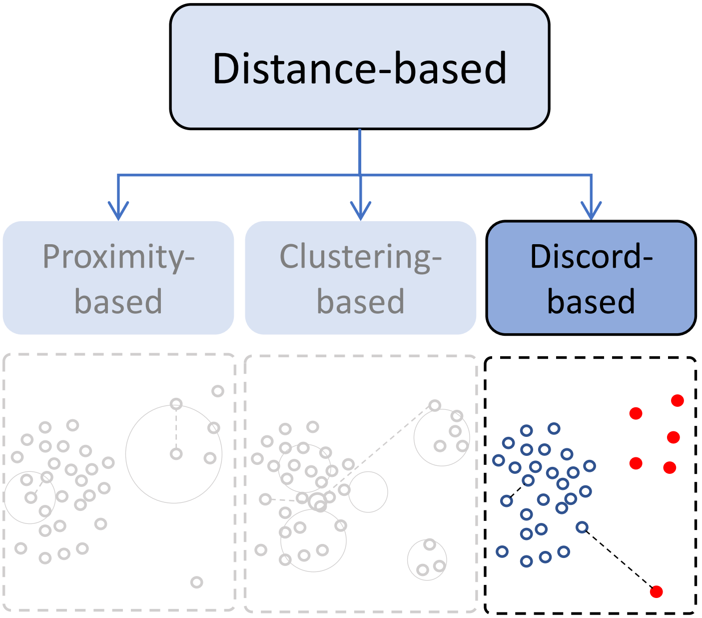
Time Series Anomaly Detection Methods

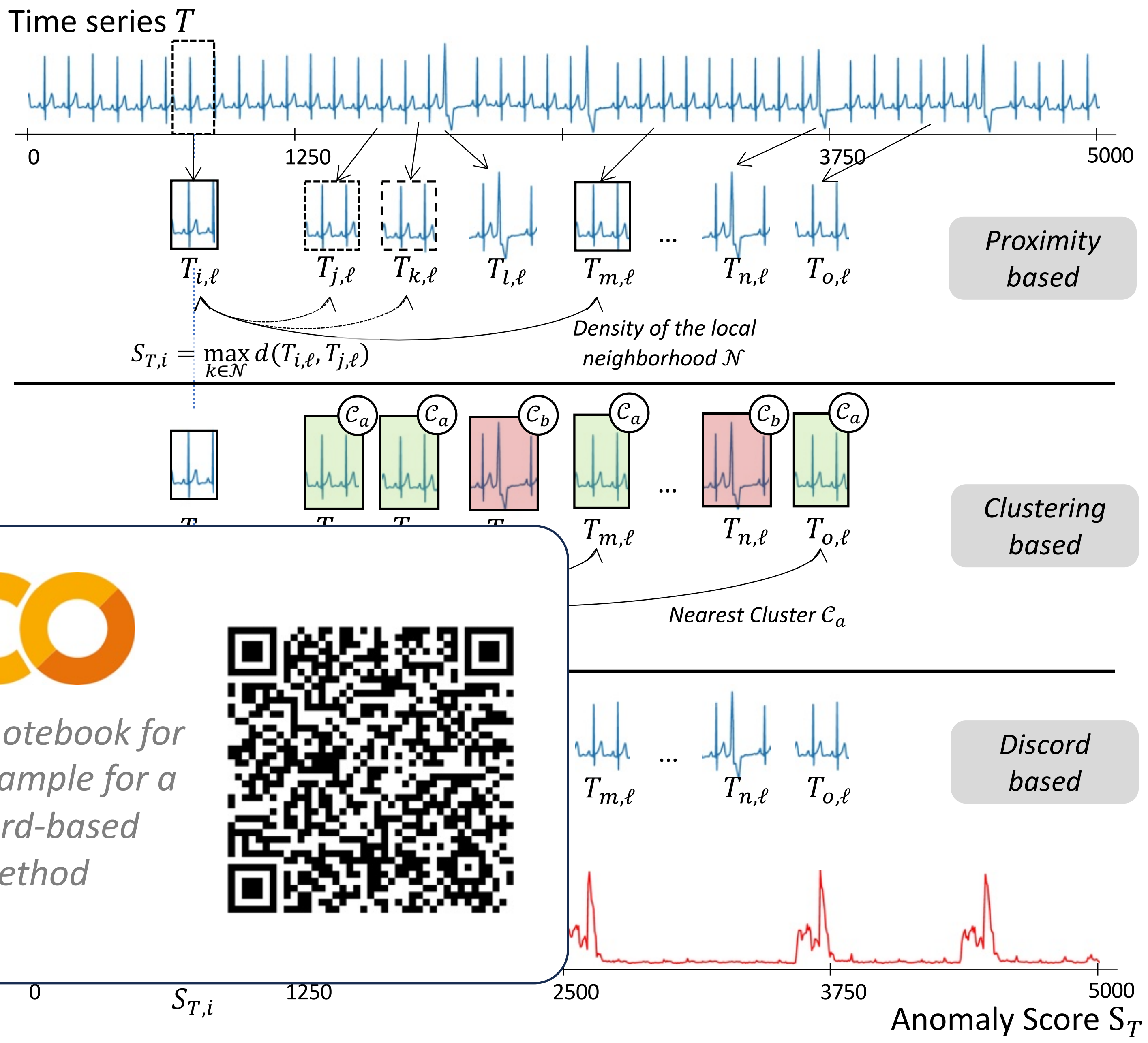
Anomaly Detection methods:
Distance-based





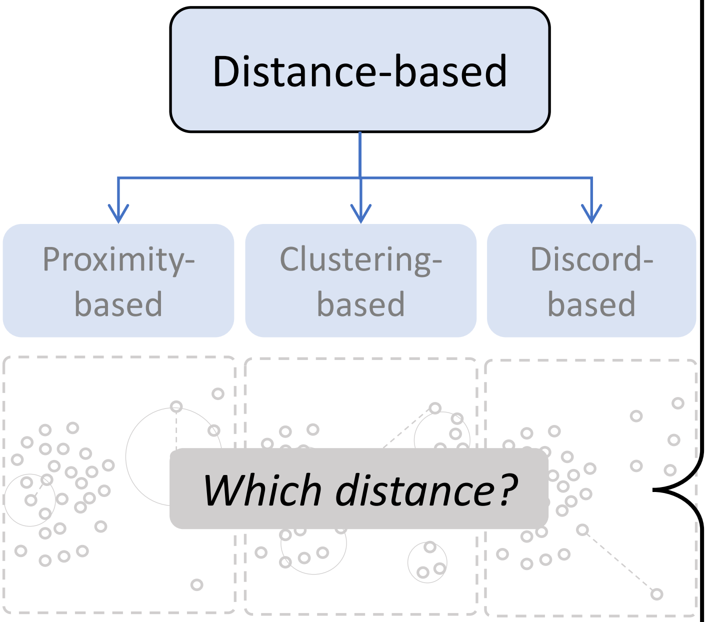
Anomaly Detection methods: *Distance-based*

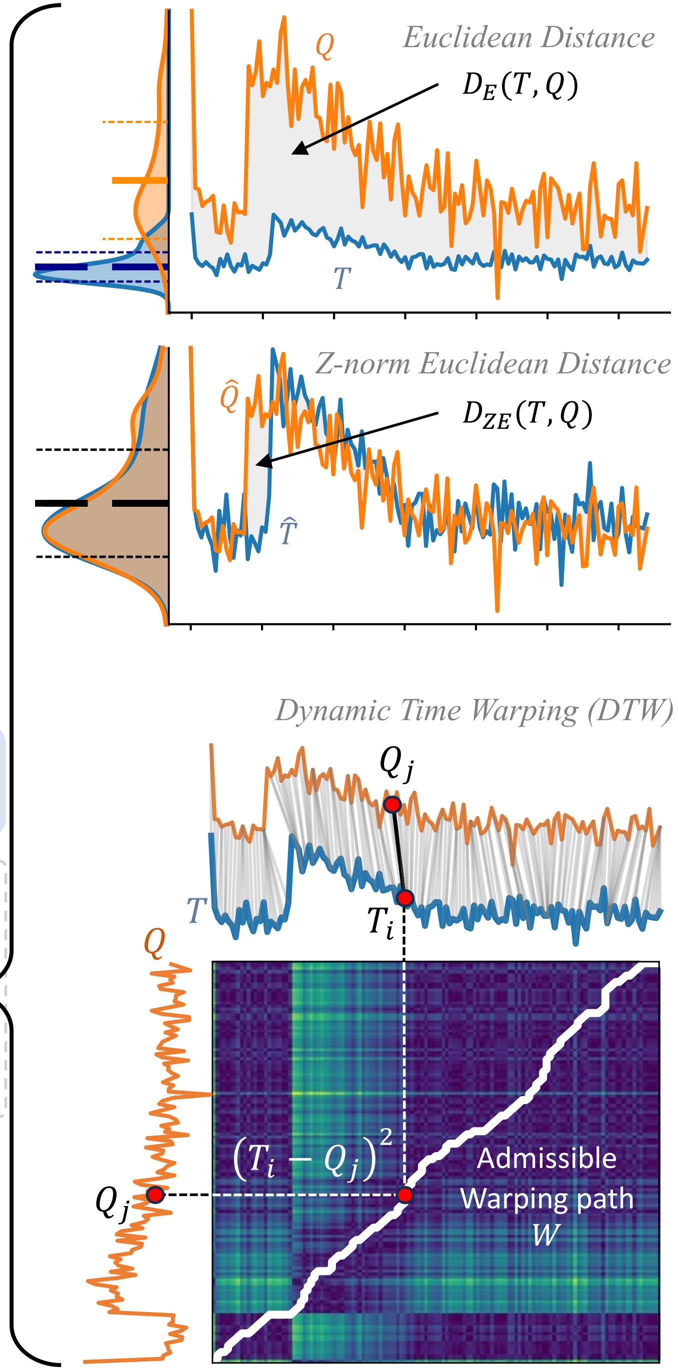




Online notebook for code example for a Discord-based method

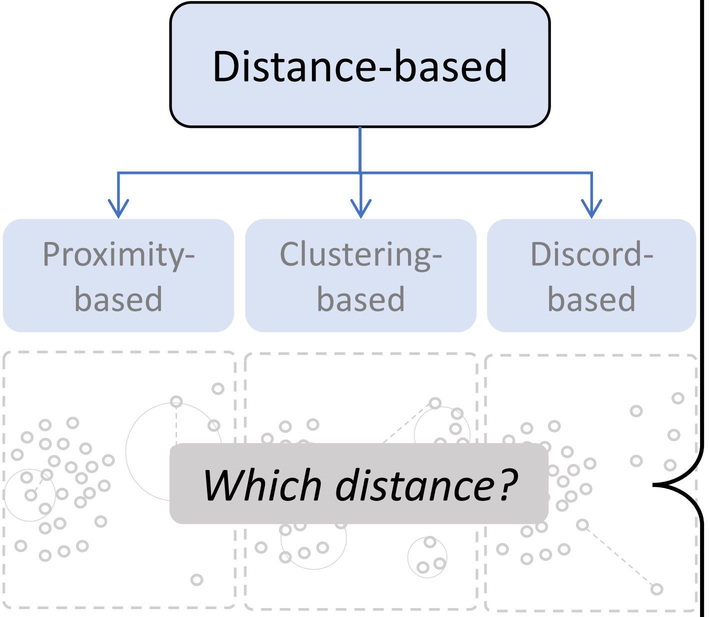
Anomaly Detection methods: *Distance-based*

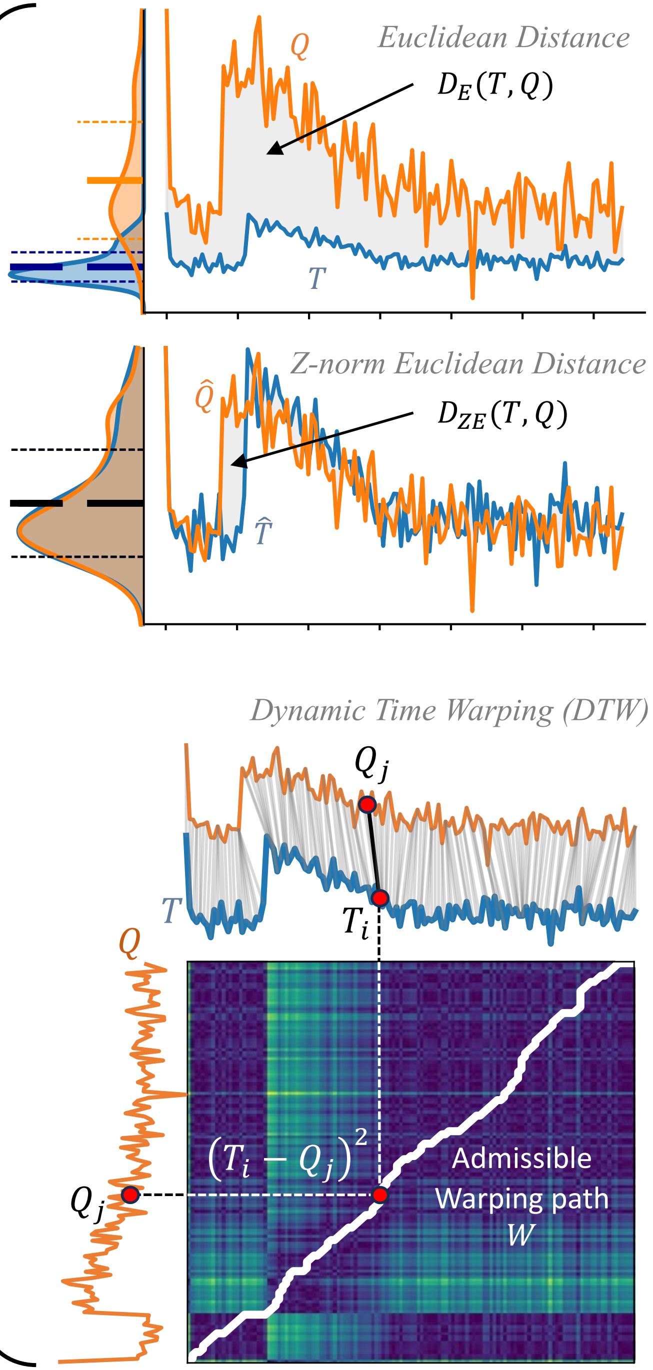




Distance	Formula	Time Complexity
Linear and Superlinear methods		
ED	$D(T,Q) = \sqrt{\sum_{i=0}^{n-1}(T_i - Q_i)^2}$	$O(n)$
z-norm ED	$D(T,Q) = \sqrt{\sum_{i=0}^{n-1}\left(\frac{T_i-\mu_T}{\sigma_T} - \frac{Q_i-\mu_Q}{\sigma_Q}\right)^2}$	$O(n)$
Pearson	$D(T,Q) = \sum_{i=0}^{n-1}\frac{(T_i - Q_i)^2}{Q_i}$	$O(n)$
Manhattan	$D(T,Q) = \sum_{i=0}^{n-1}\lvert T_i - Q_i\rvert$	$O(n)$
Chebyshev	$D(T,Q) = \max_{0\leq i<n}\lvert T_i - Q_i\rvert$	$O(n)$
Cosine	$D(T,Q) = 1 - \frac{\sum_i T_i Q_i}{\sqrt{\sum_i T_i^2}\sqrt{\sum_i Q_i^2}}$	$O(n)$
SBD [99]	$D(T,Q) = 1 - \max_{\tau}\frac{\sum_{i=0}^{n-1}(T_i - \bar{T})(Q_{i+\tau} - \bar{Q})}{\sqrt{\sum_{i=0}^{n-1}(T_i - \bar{T})^2}\sqrt{\sum_{i=0}^{n-1}(Q_{i+\tau} - \bar{Q})^2}}$	$O(n \log n)$
Quadratic methods ($D(T,Q) = \sqrt{D(T_{n-1}, Q_{n-1})}$)		
DTW	$D(T_i,Q_j) = (T_i - Q_j)^2 + \min\begin{cases} D(T_{i-1},Q_j), \\ D(T_i,Q_{j-1}), \\ D(T_{i-1},Q_{j-1}) \end{cases}$	$O(n^2)$
LCSS [132]	$D(T_i,Q_j) = \begin{cases} 0 & i=0 \text{ or } j=0, \\ D(T_{i-1},Q_{j-1})+1 & \text{if } \lvert T_i - Q_j\rvert \leq \epsilon, \\ \max\begin{cases} D(T_{i-1},Q_j), \\ D(T_i,Q_{j-1}) \end{cases} & \text{otherwise.} \end{cases}$	$O(n^2)$
ERP [30]	$D(T_i,Q_j) = \begin{cases} \sum_{k=0}^{i}\lvert T_k - g\rvert & j=0, \\ \sum_{k=0}^{j}\lvert Q_k - g\rvert & i=0, \\ \min\begin{cases} D(T_{i-1},Q_{j-1}) + \lvert T_i - Q_j\rvert, \\ D(T_{i-1},Q_j) + \lvert T_i - g\rvert, \\ D(T_i,Q_{j-1}) + \lvert Q_j - g\rvert \end{cases} & \text{otherwise.} \end{cases}$	$O(n^2)$
MSM [121]	$D(T_i,Q_j) = \min\begin{cases} D(T_{i-1},Q_{j-1}) + \lvert T_i - Q_j\rvert, \\ D(T_{i-1},Q_j) + C(T_i,T_{i-1},Q_j), \\ D(T_i,Q_{j-1}) + C(Q_j,T_i,Q_{j-1}) \end{cases}$ with $C(x,y,z) = \begin{cases} c & \text{if } x \in [y,z] \text{ or } x \in [z,y], \\ c + \min\{\lvert x-y\rvert, \lvert x-z\rvert\} & \text{otherwise.} \end{cases}$	$O(n^2)$
TWED [83]	$D(T_i,Q_j) = \begin{cases} \sum_{k=1}^{i}\lvert T_k - T_{k-1}\rvert + \lambda i & j=0, \\ \sum_{k=1}^{j}\lvert Q_k - Q_{k-1}\rvert + \lambda j & i=0, \\ \min\begin{cases} D(T_{i-1},Q_{j-1}) + \lvert T_i - Q_j\rvert + \nu\lvert i-j\rvert, \\ D(T_{i-1},Q_j) + \lvert T_i - T_{i-1}\rvert + \lambda + \nu, \\ D(T_i,Q_{j-1}) + \lvert Q_j - Q_{j-1}\rvert + \lambda + \nu \end{cases} & \text{otherwise.} \end{cases}$	$O(n^2)$

Anomaly Detection methods:
Distance-based

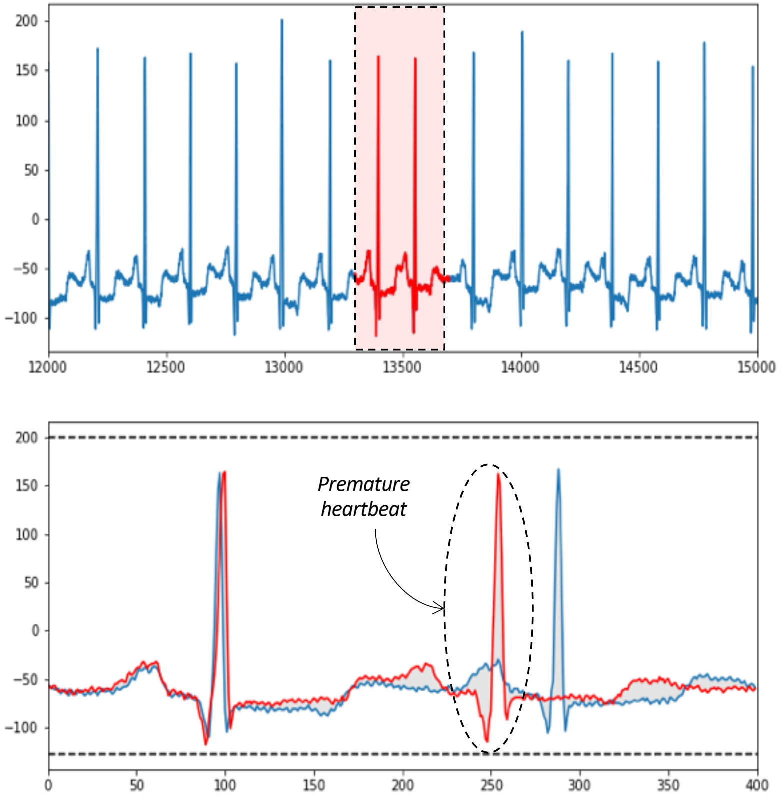




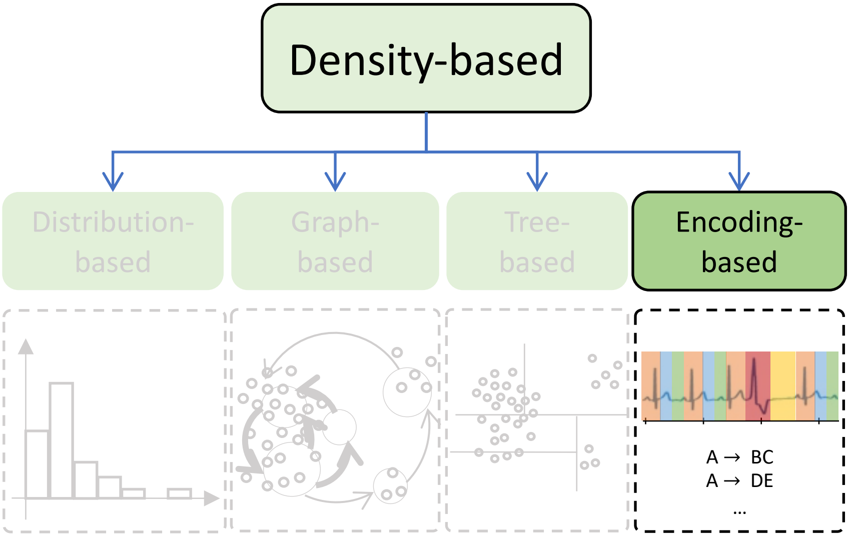
Distance	Formula	Time Complexity
Linear and Superlinear methods		
ED	$D(T,Q) = \sqrt{\sum_{i=0}^{n-1}(T_i - Q_i)^2}$	$O(n)$
z-norm ED	$D(T,$	$O(n)$
Pearson	$D(T,$	$O(n)$
Manhattan	$D(T,$	$O(n)$
Chebyshev	$D(T,Q) = \max_{0 \le i < n} \lvert T_i - Q_i \rvert$	$O(n)$
Cosine		
SBD [99]		)
DTW		
LCSS [132		
ERP [30]		
MSM [12]		
TWED [8	$D(T_i, Q_{j-1}) + \lvert Q_j - Q_{j-1} \rvert + \lambda + \nu$	

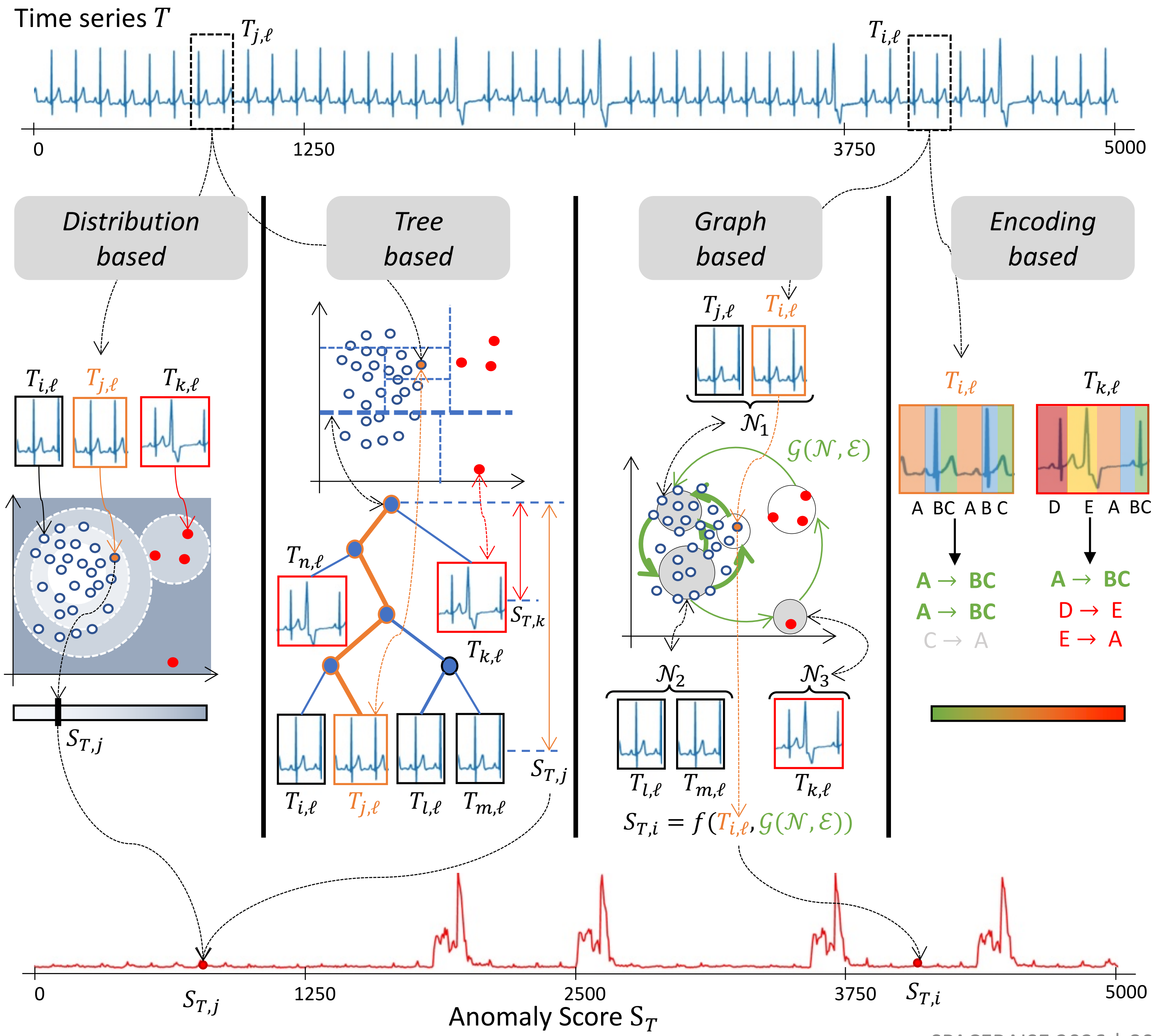
Which distance is the most appropriate for time series anomaly detection?

(a) Data series example: Snippet of an electrocardiogram (in blue: normal heartbeats, in red: premature heartbeat)

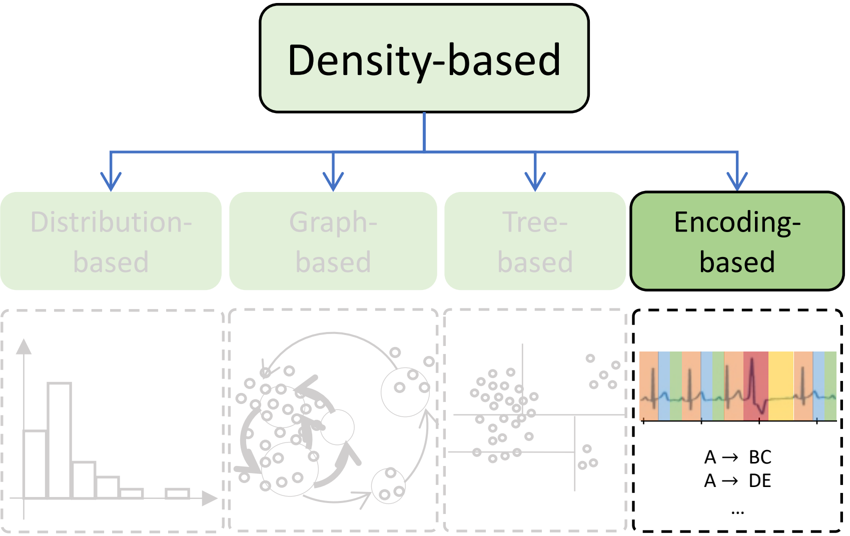


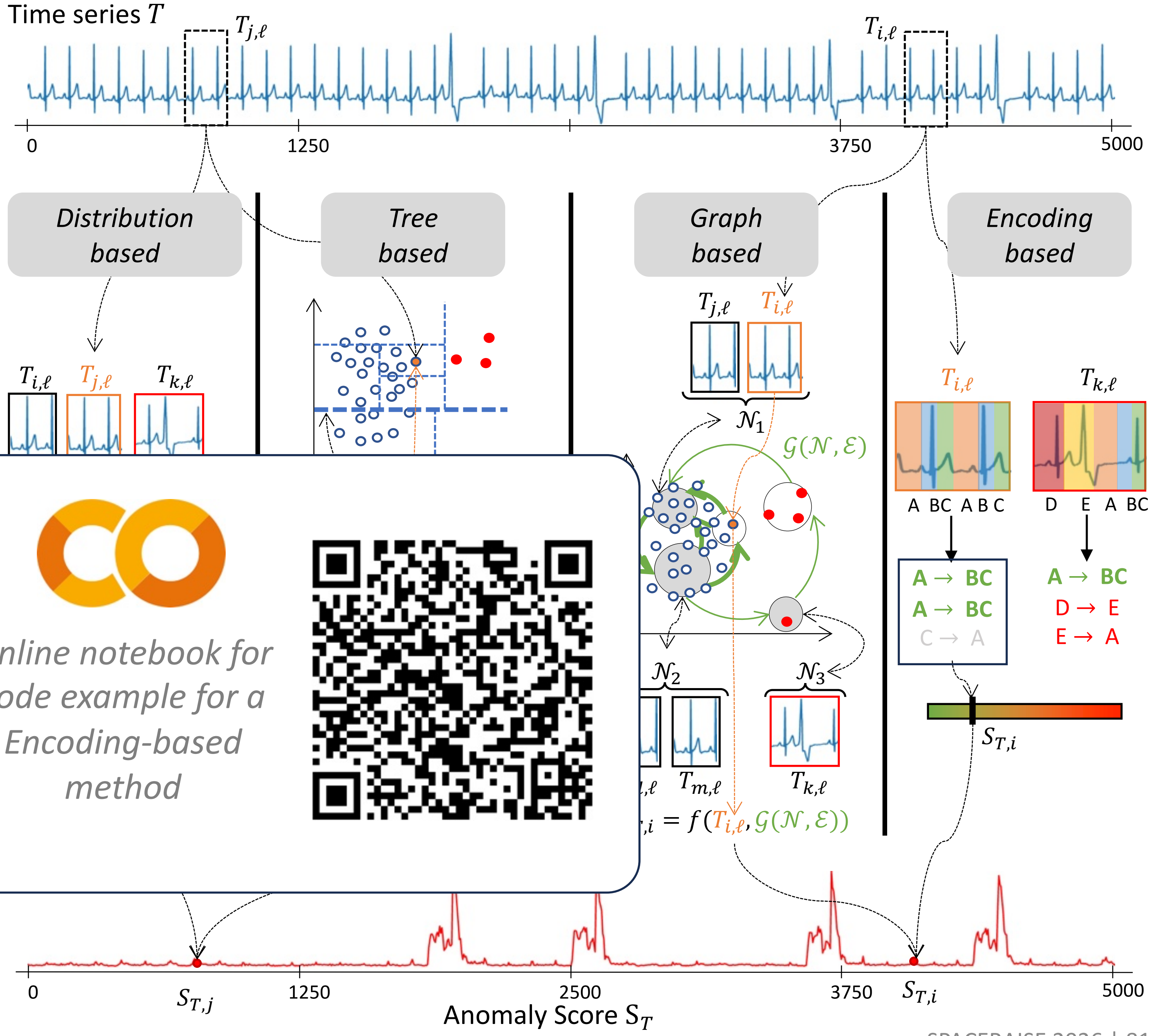
Anomaly Detection methods: *Density-based*





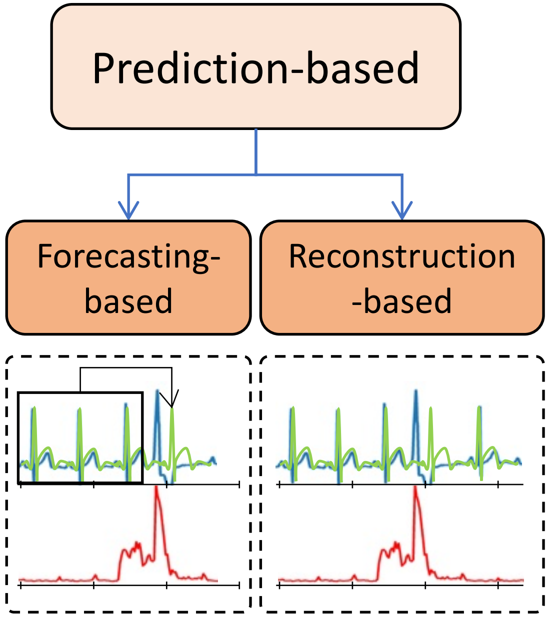
Anomaly Detection methods: *Density-based*

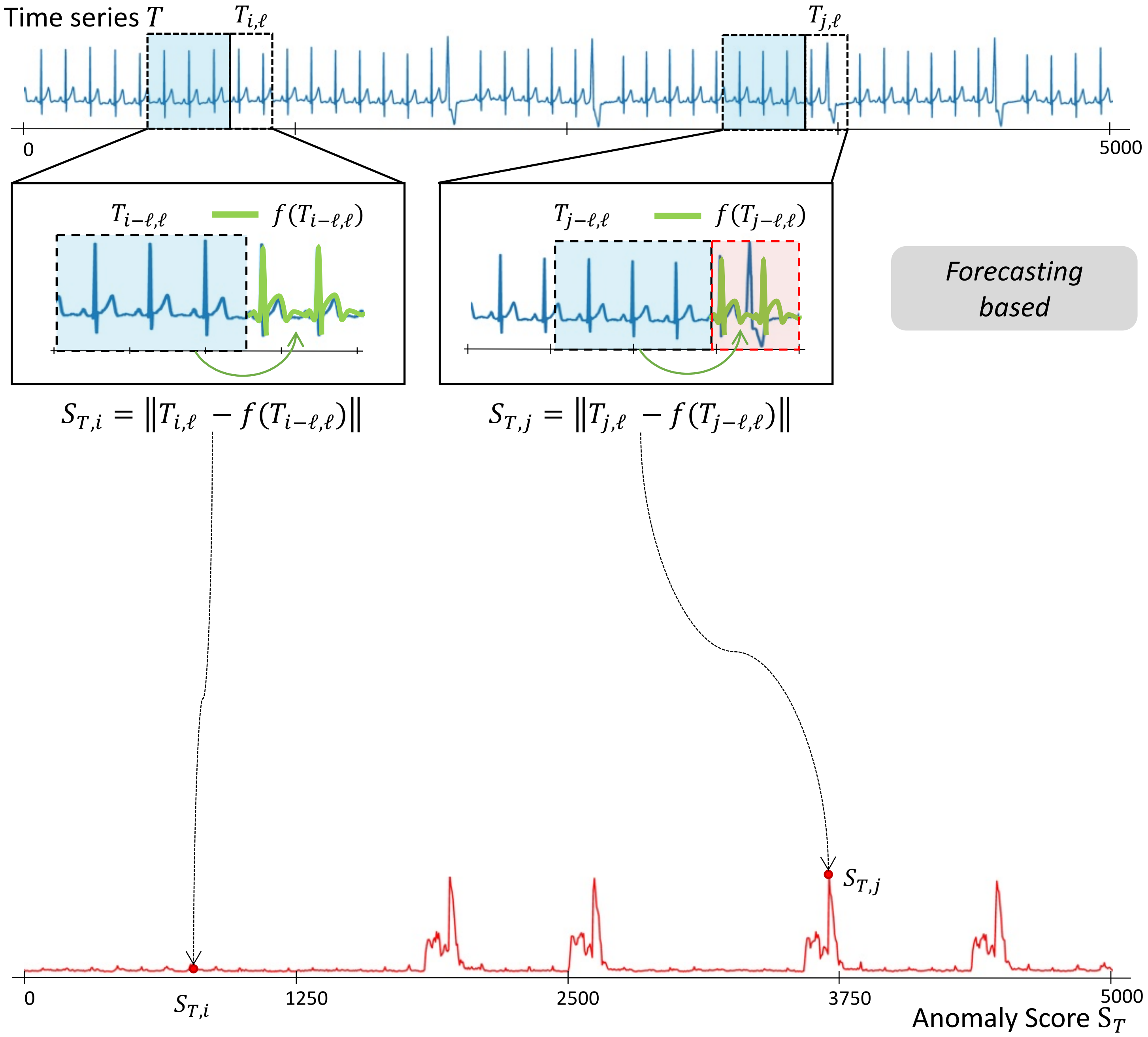




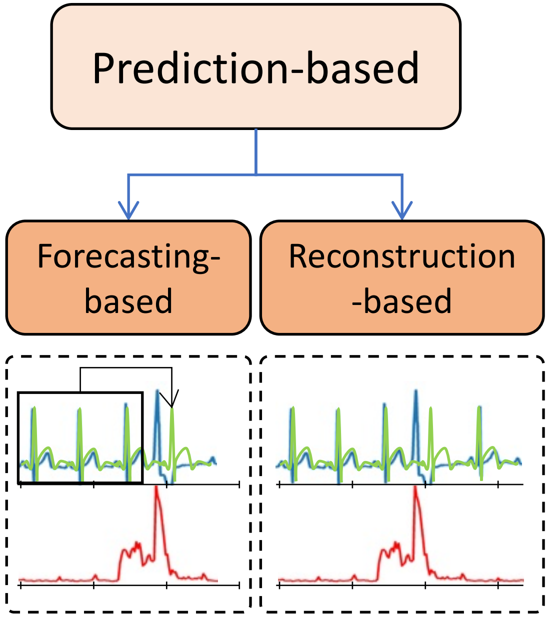
Online notebook for code example for a Encoding-based method

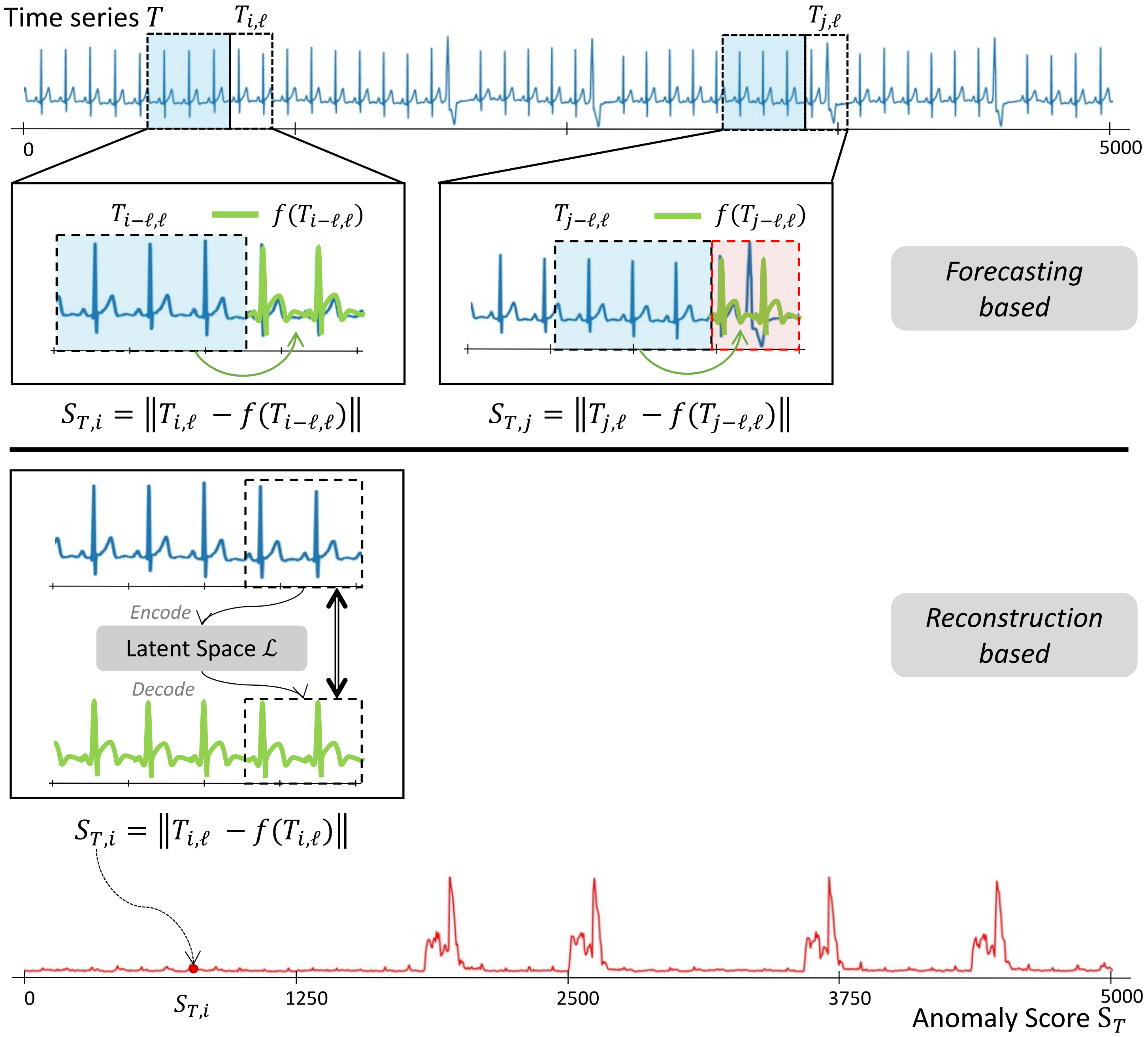
Anomaly Detection methods: *Prediction-based*



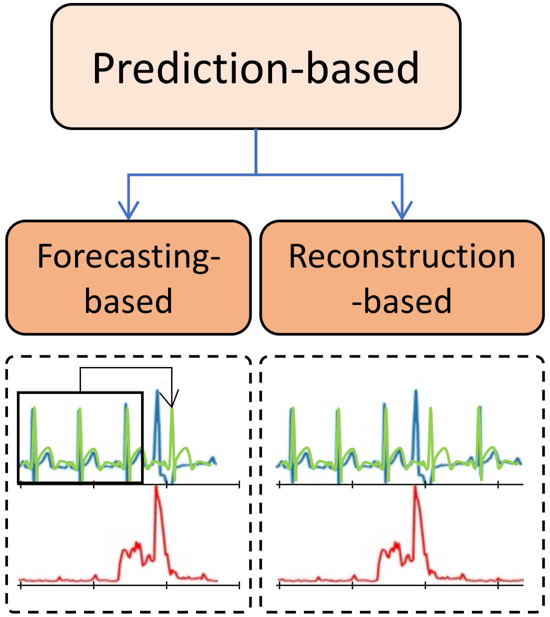


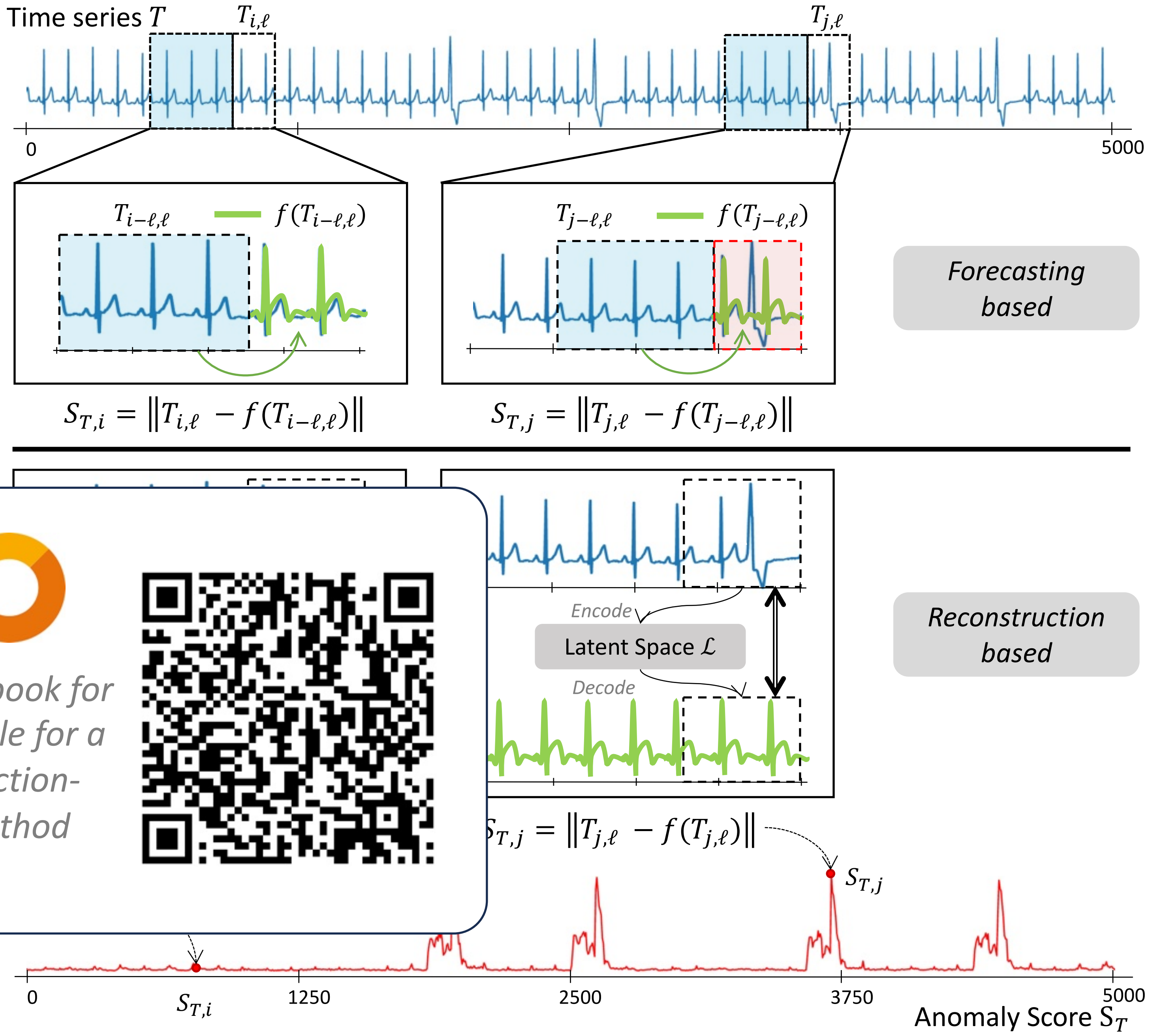
Anomaly Detection methods: *Prediction-based*



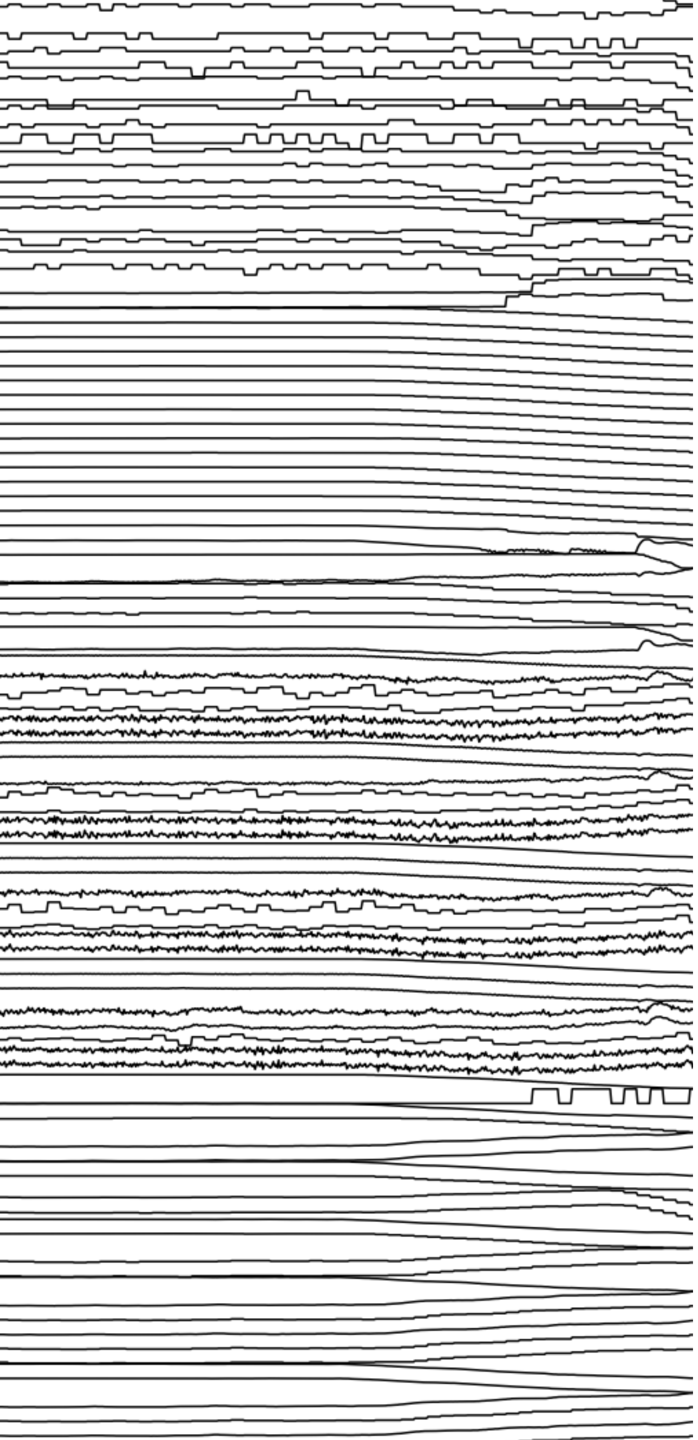


Anomaly Detection methods: *Prediction-based*





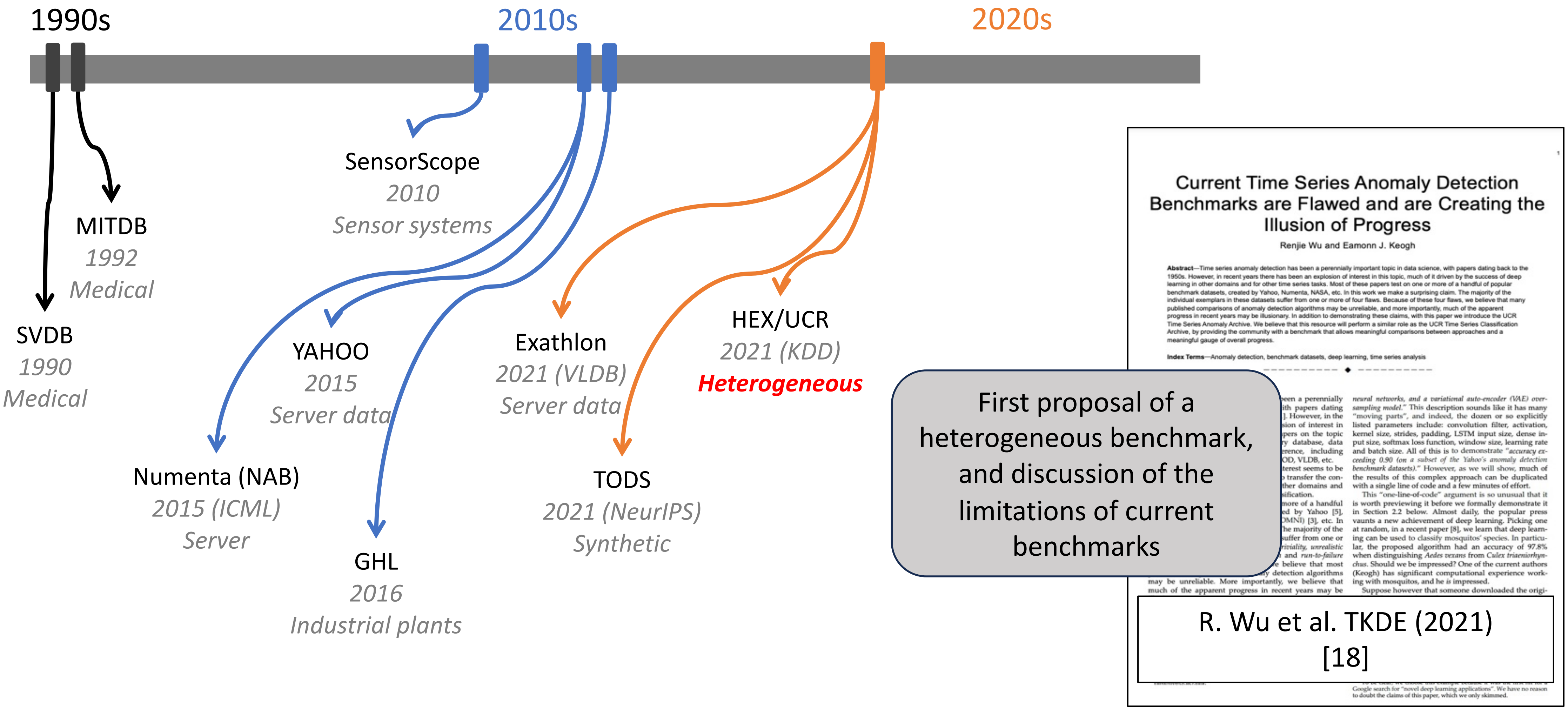
Online notebook for code example for a Reconstruction-based method



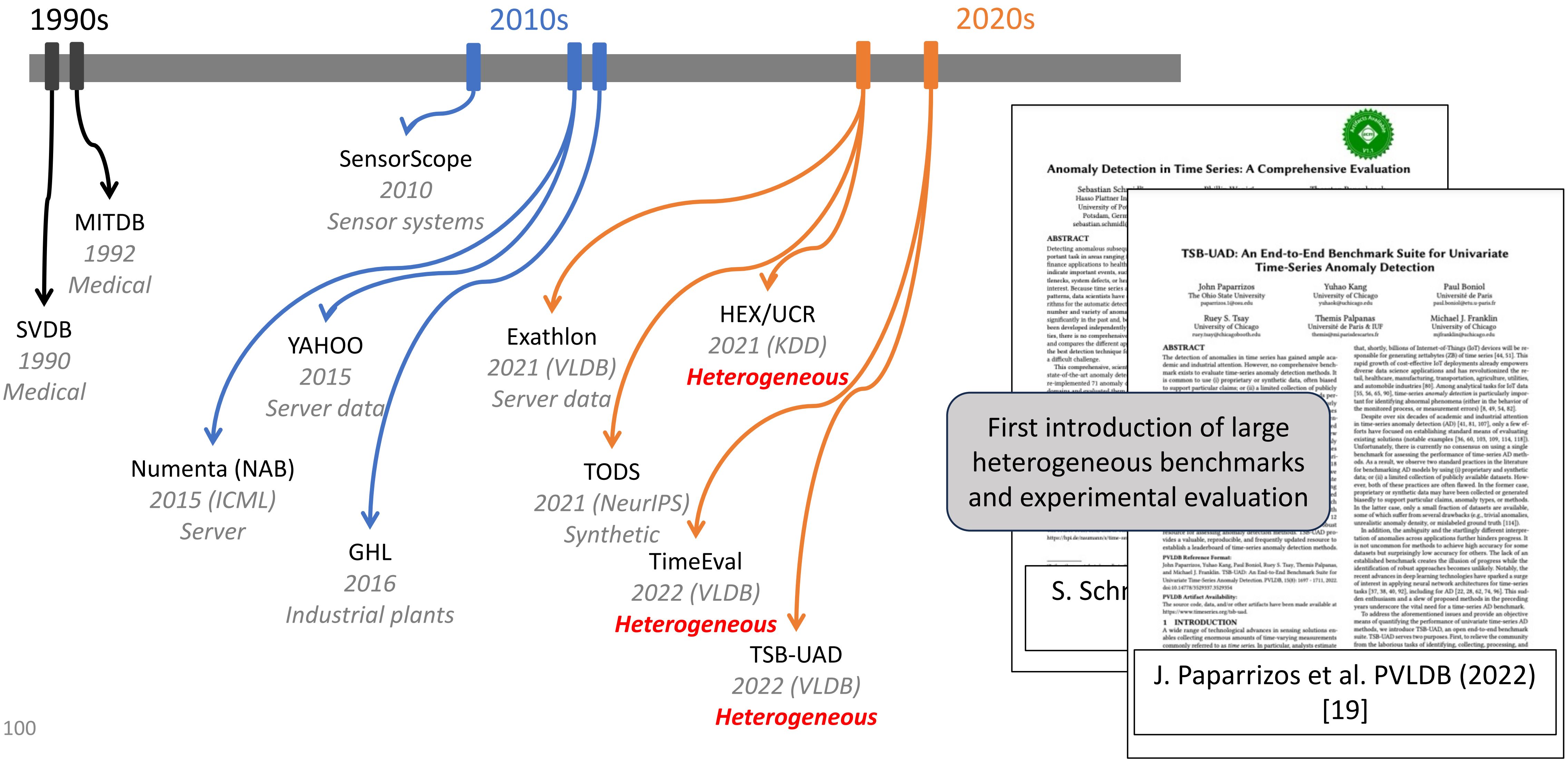
Part III.

Benchmark, Evaluation Measures, and Automatic Solutions

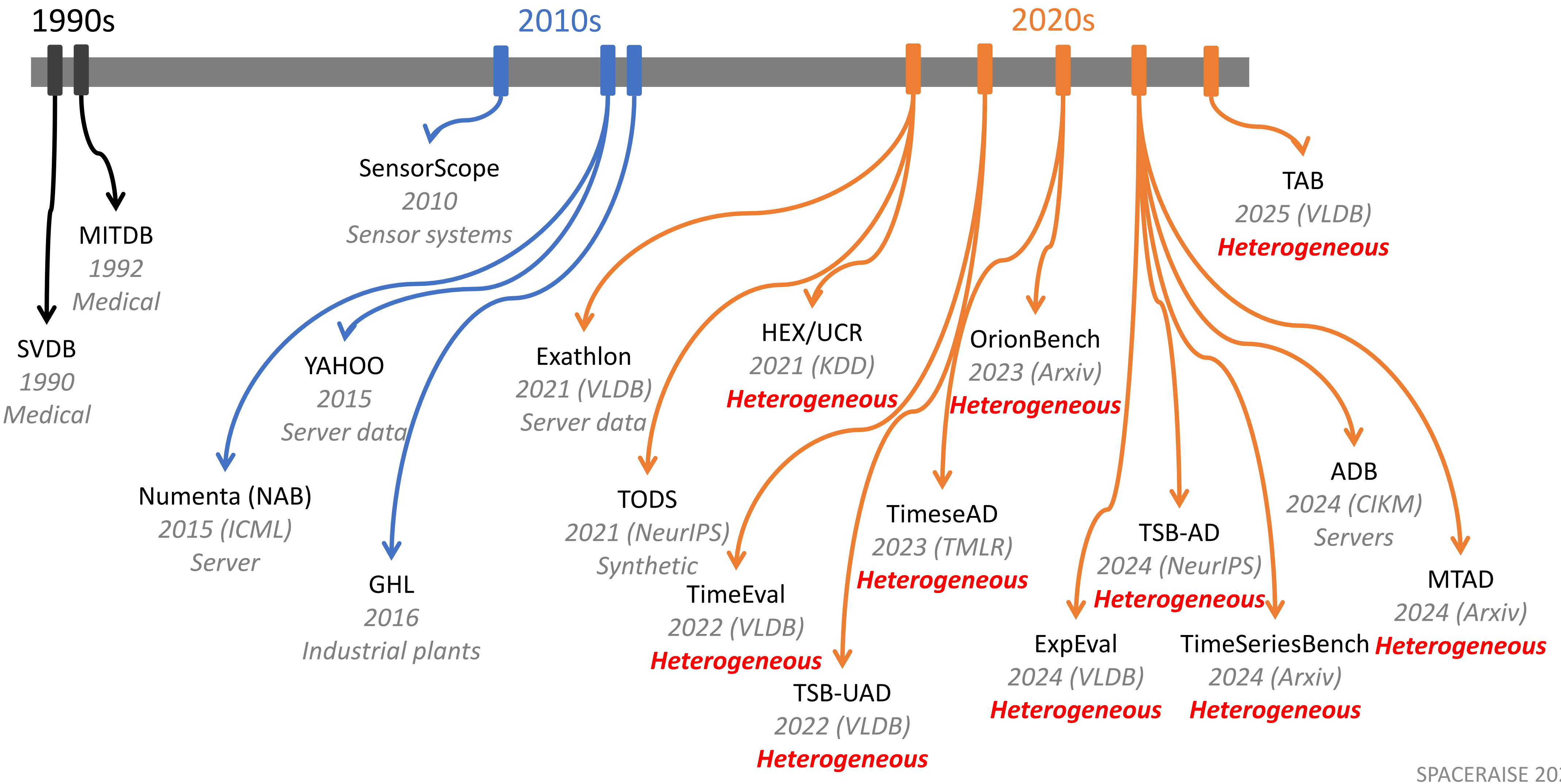
Anomaly Detection methods: *Existing benchmark*



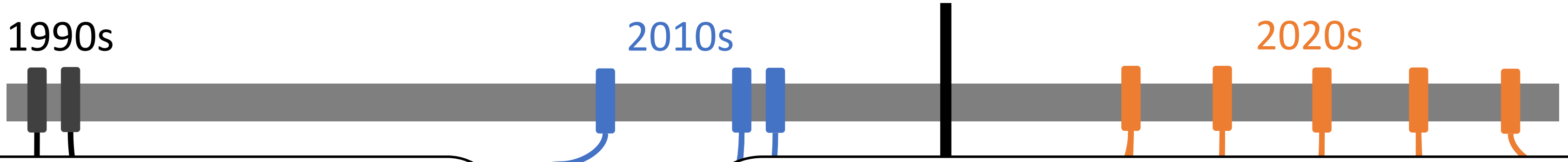
Anomaly Detection methods: *Existing benchmark*



Anomaly Detection methods: *Existing benchmark*

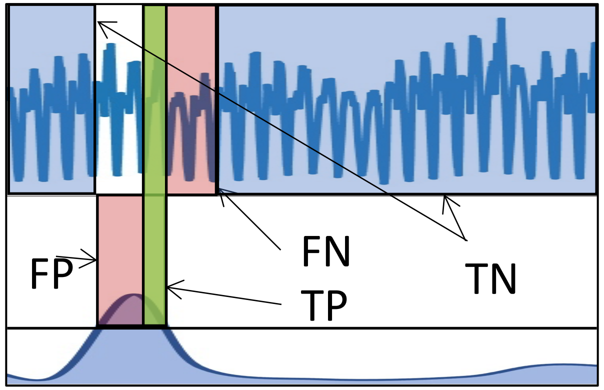


Anomaly Detection methods: *Existing benchmark*



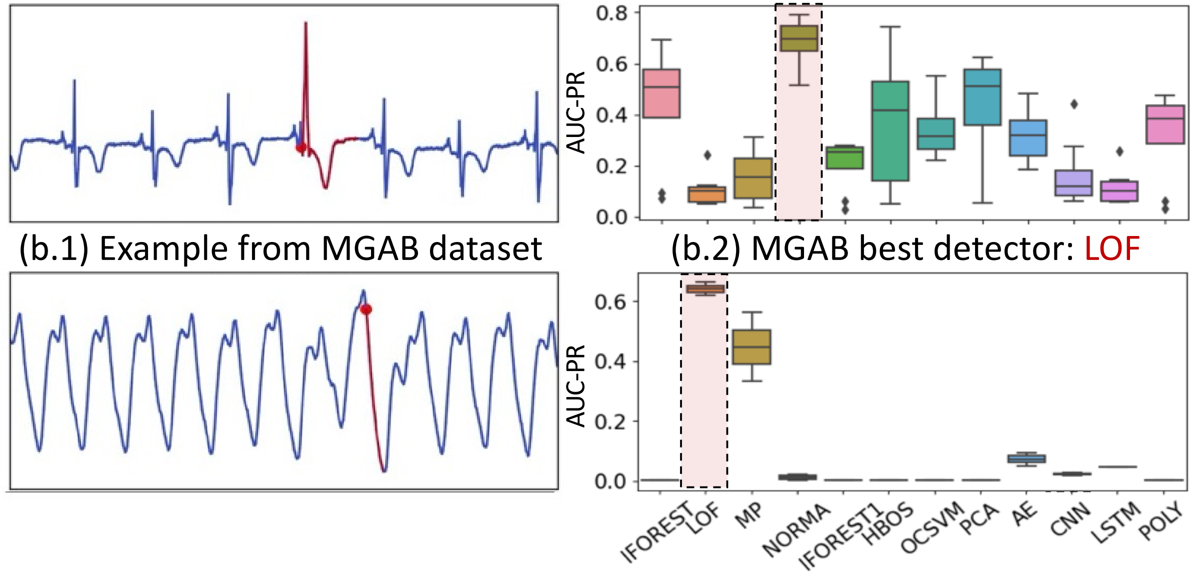
How should we **evaluate** the **accuracy** of time series anomaly detection?

Traditional point-wise and threshold-dependent measures are over-penalizing

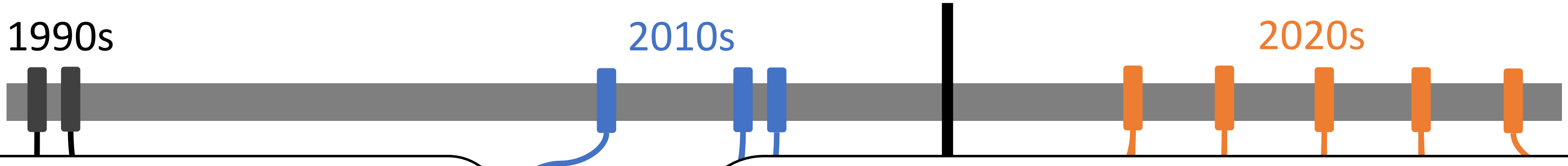


Can we **rely** on one **unique** method?

Observations from recent heterogeneous benchmarks (i.e., TSB-UAD, TSB-AD, TimeEval, and TAB) confirm that there are significant differences in performance across datasets

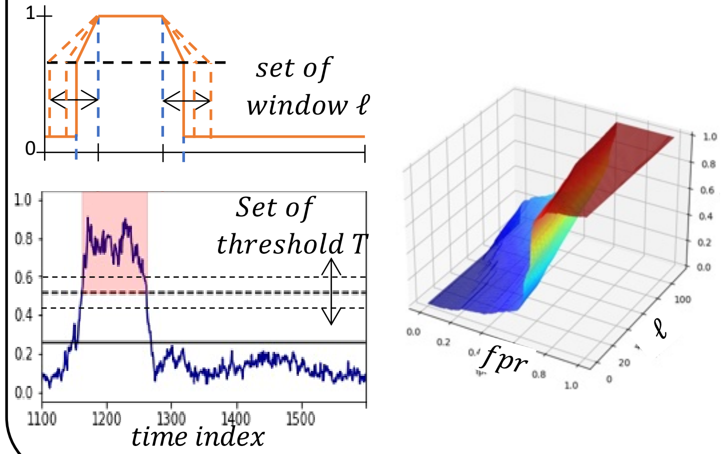


Anomaly Detection methods: *Existing benchmark*



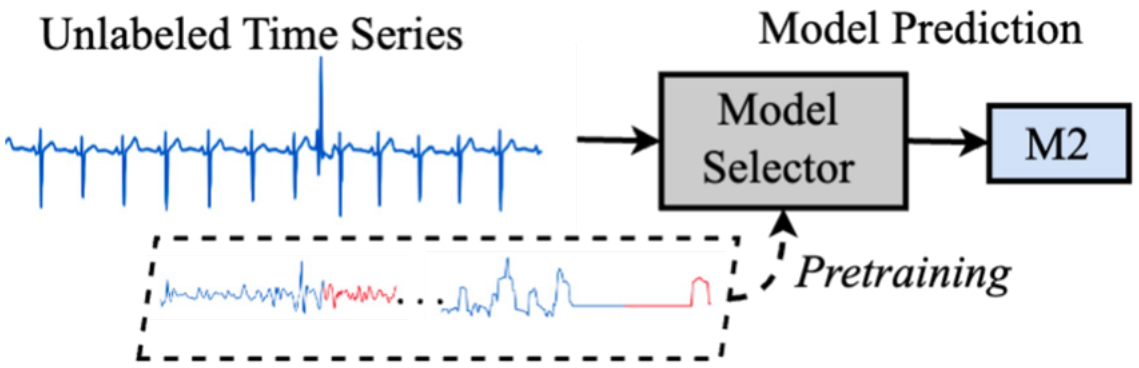
How should we **evaluate** the **accuracy** of time series anomaly detection?

Recent time series specific measures (such as VUS) have been introduced

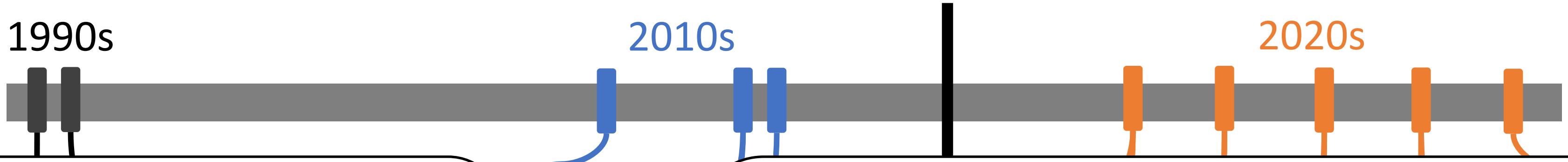


Can we **rely** on one **unique** method?

Model Selection and ensembling significantly outperform each individual detectors



Anomaly Detection methods: *Existing benchmark*



How should we **evaluate** the **accuracy** of time series anomaly detection?

Recent time series specific measures (such as VUS) have been

More info on previous lectures:

https://boniolp.github.io/TwinODIS

(Part III)

Can we **rely** on one **unique** method?

Model Selection and ensembling significantly outperform each individual detectors

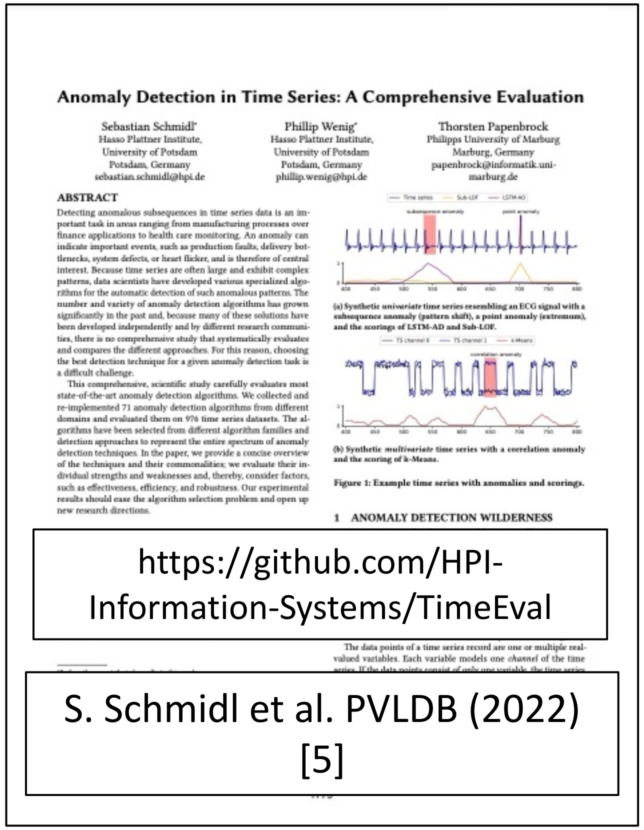
More info on previous lectures:

https://boniolp.github.io/TwinODIS

(Part III)

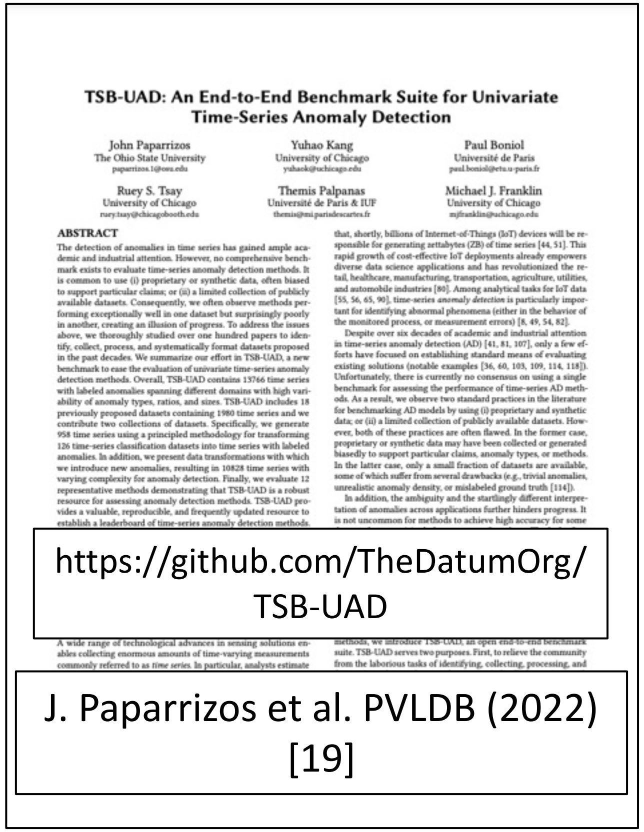
Conclusion

If you are interested in anomaly detection in time series…



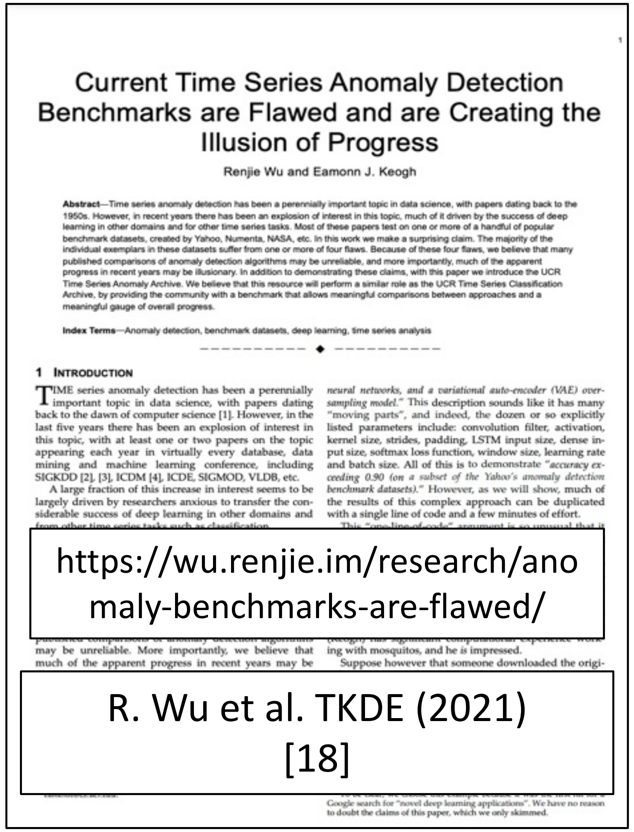
https://github.com/HPI-Information-Systems/TimeEval

S. Schmidl et al. PVLDB (2022) [5]



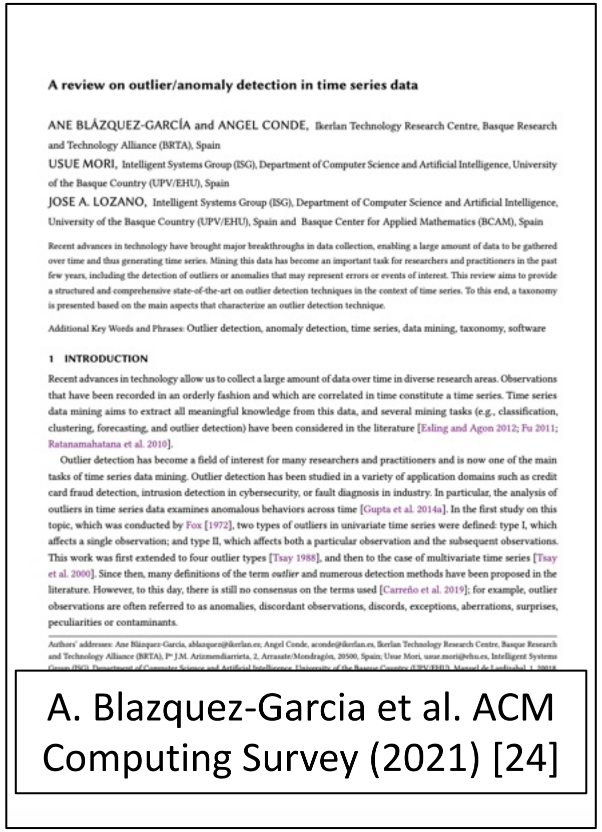
https://github.com/TheDatumOrg/TSB-UAD

J. Paparrizos et al. PVLDB (2022) [19]



https://wu.renjie.im/research/anomaly-benchmarks-are-flawed/

R. Wu et al. TKDE (2021) [18]



A. Blazquez-Garcia et al. ACM Computing Survey (2021) [24]

Conclusion

Time Series Anomaly Detection

Paul Boniol, Qinghua Liu, John Paparrizos, and Themis Palpanas.

Video *(EDBT 2023 Tutorial)*	Slides *(VLDB 2024 Tutorial)*	Short Survey *(KDD 2025)*	Survey *(2025 preprint)*
https://www.youtube.com/watch?v=96869qimXAA&t=1s	https://drive.google.com/file/d/1Vyz6H0E16IpbVZXgtiZVnU9Ie8zAJaog/view	https://dl.acm.org/doi/10.1145/3711896.3736565	https://arxiv.org/pdf/2412.20512

S. Schmidl et al. PVLDB (2022) [5]

J. Paparrizos et al. PVLDB (2022) [19]

R. Wu et al. TKDE (2021) [18]

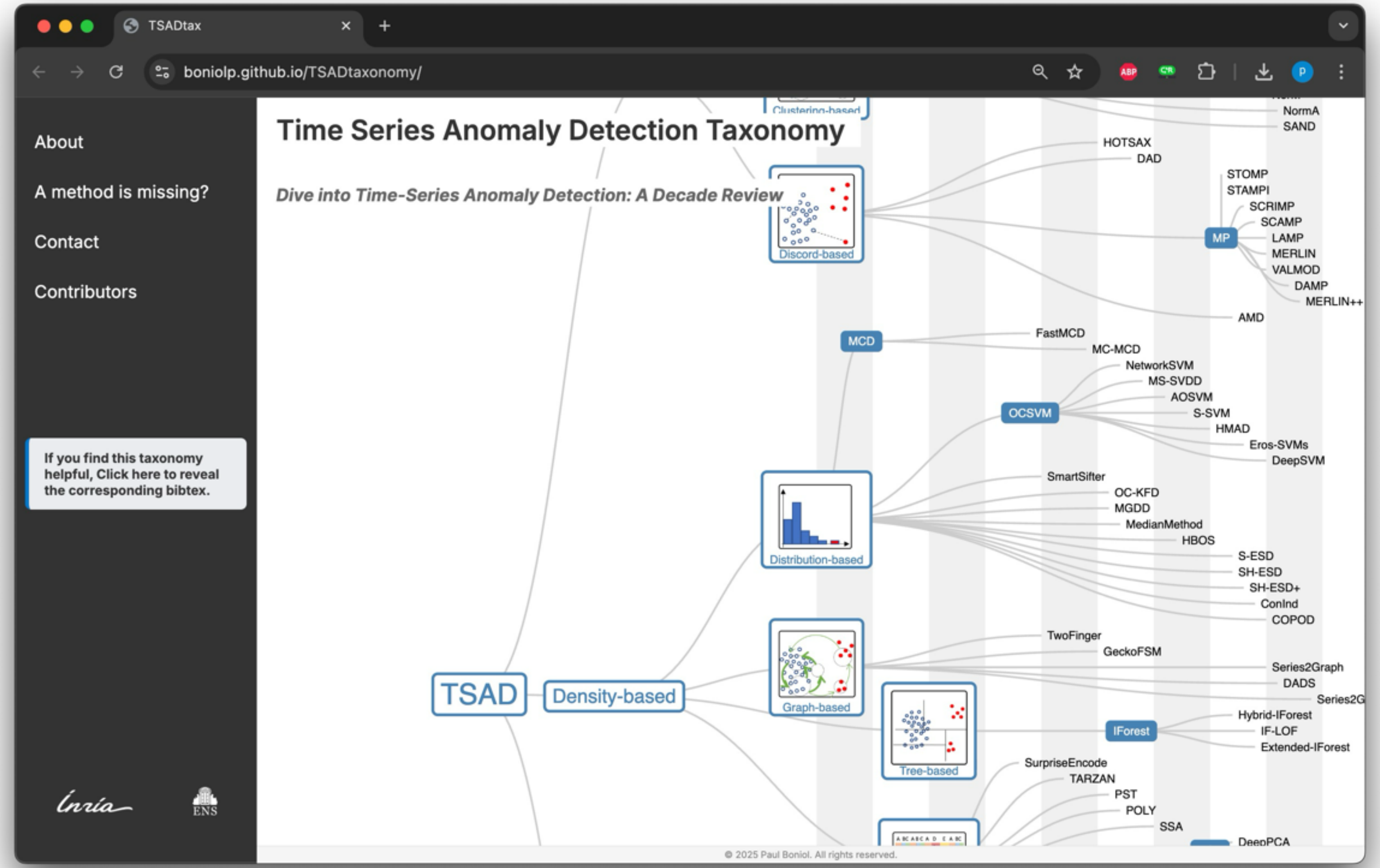
A. Blazquez-Garcia et al. ACM Computing Survey (2021) [24]

Conclusion

Online interactive Taxonomy

https://boniolp.github.io/TSADtaxonomy/

https://github.com/boniolp/TSADtaxonomy



Conclusion

TimeEval

PyPI v1.1.4, Python 3.7 – 3.9

`Pip install TimeEval`

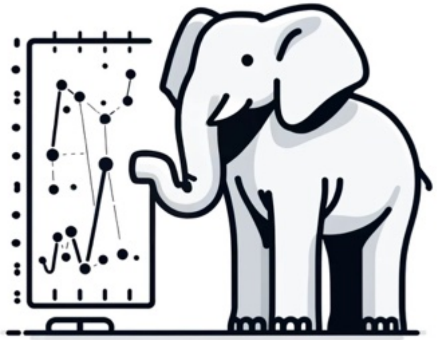
github.com/TimeEval/TimeEval

PyPI v0.0.3, Python 3.8 – 3.12

`Pip install tsb-uad`

github.com/TheDatumOrg/TSB-UAD

TSB-AD



No PyPI, Python 3.9

`Pip install TSB-AD`

github.com/TheDatumOrg/TSB-AD

Conclusion

PyPI v0.0.3, Python 3.8 – 3.12

```
Pip install tsb-uad
```

github.com/TheDatumOrg/TSB-UAD

```python
import os
import numpy as np
import pandas as pd
from TSB_UAD.models.iforest import IForest
from TSB_UAD.models.feature import Window
from TSB_UAD.utils.slidingWindows import find_length
from TSB_UAD.vus.metrics import get_metrics

df = pd.read_csv('data/benchmark/ECG/MBA_ECG805_data.out', header=None).to_numpy()
data = df[:, 0].astype(float)
label = df[:, 1]

slidingWindow = find_length(data)
X_data = Window(window = slidingWindow).convert(data).to_numpy()

clf = IForest(n_jobs=1)
clf.fit(X_data)
score = clf.decision_scores_

score = MinMaxScaler(feature_range=(0,1)).fit_transform(score.reshape(-1,1)).ravel()
score = np.array([score[0]]*math.ceil((slidingWindow-1)/2) + list(score) + [score[-1]]*((s


results = get_metrics(score, label, metric="all", slidingWindow=slidingWindow)
for metric in results.keys():
    print(metric, ':', results[metric])
```

Conclusion

PyPI v0.0.3, Python 3.8 – 3.12

```
Pip install tsb-uad
```

github.com/TheDatumOrg/TSB-UAD

```python
import os
import numpy as np
import pandas as pd
from TSB_UAD.models.iforest import IForest
from TSB_UAD.models.feature import Window
from TSB_UAD.utils.slidingWindows import find_length
from TSB_UAD.vus.metrics import get_metrics

df = pd.read_csv('data/benchmark/ECG/MBA_ECG805_data.out', header=None).to_numpy()
data = df[:, 0].astype(float)
label = df[:, 1]

slidingWindow = find_length(data)
X_data = Window(window = slidingWindow).convert(data).to_numpy()

clf = IForest(n_jobs=1)
clf.fit(X_data)
score = clf.decision_scores_

score = MinMaxScaler(feature_range=(0,1)).fit_transform(score.reshape(-1,1)).ravel()
score = np.array([score[0]]*math.ceil((slidingWindow-1)/2) + list(score) + [score[-1]]*((s


results = get_m
for metric in r
    print(metri
```

```
AUC_ROC : 0.9216216369841076
AUC_PR : 0.6608577550833885
Precision : 0.7342093339374717
Recall : 0.4010891089108911
F : 0.5187770129662238
Precision_at_k : 0.4010891089108911
Rprecision : 0.7486112853253205
Rrecall : 0.30977335423161651
RF : 0.438214653167952
R_AUC_ROC : 0.989123018780308
R_AUC_PR : 0.9435238401582703
VUS_ROC : 0.9734357459251715
VUS_PR : 0.8858037295594041
Affiliation_Precision : 0.9630674176380548
Affiliation_Recall : 0.9809813654809071
```

Conclusion

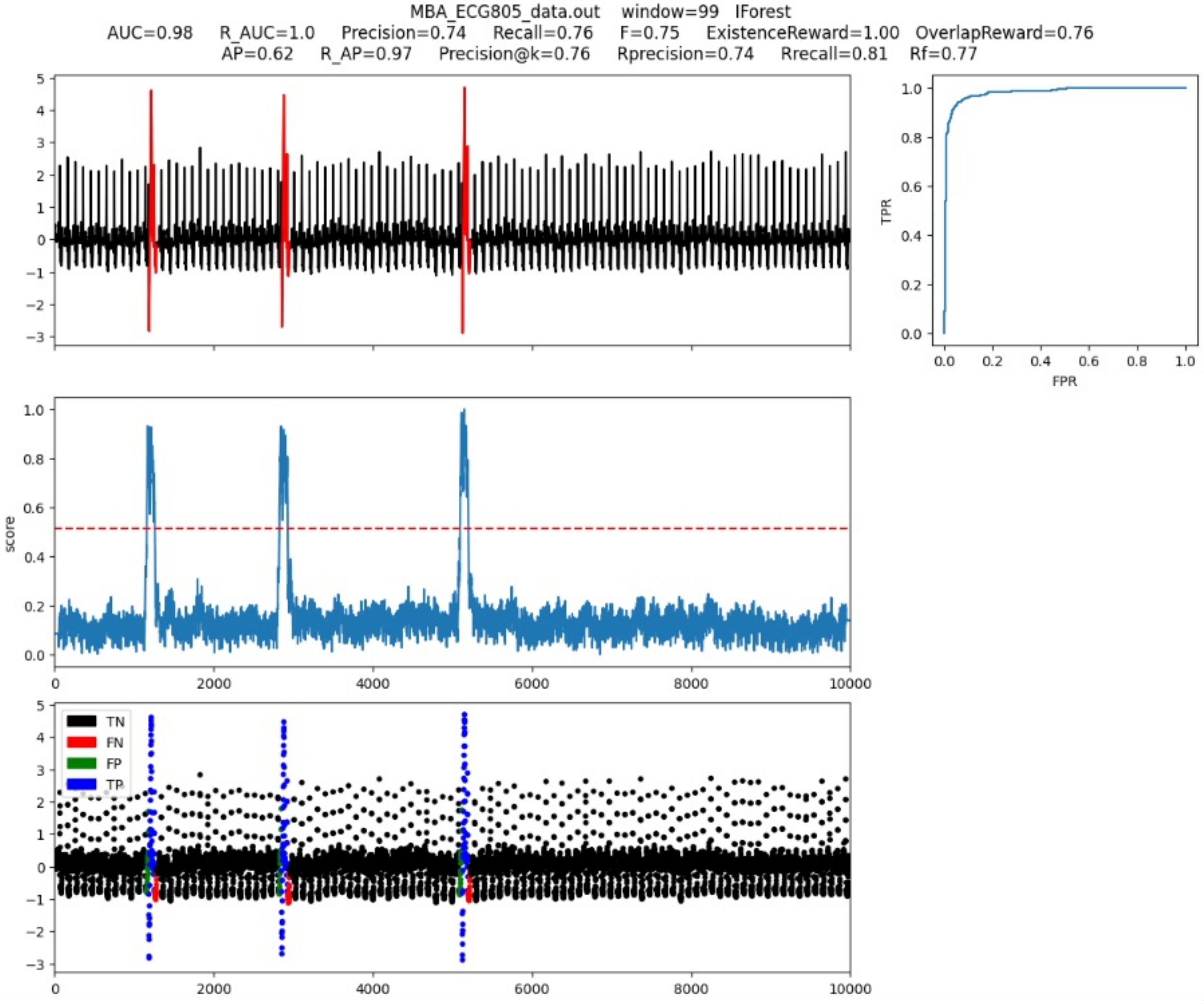
PyPI v0.0.3, Python 3.8 – 3.12

```
Pip install tsb-uad
```

github.com/TheDatumOrg/TSB-UAD

```
modelName='IForest'
clf = IForest(n_jobs=1)
x = X_data
clf.fit(x)
score = clf.decision_scores_
score = MinMaxScaler(feature_range=(0,1)).fit_transform(score.reshape(-1,1)).ravel()
score = np.array([score[0]]*math.ceil((slidingWindow-1)/2) + list(score) + [score[-1]]*((slidingWindow-1)//2))

plotFig(data, label, score, slidingWindow, fileName=name, modelName=modelName)
```



Thanks for attending!

paul.boniol@inria.fr

https://boniolp.github.io/