

# **CONTEXT-AWARE ANOMALY DETECTION FOR COMPLEX AND DYNAMIC INDUSTRIAL SETTINGS**

**Louis Carpentier**

# LET'S MEET THE TEAM



**Ir. Carpentier Louis**  
KU Leuven  
Campus Brugge



**Prof. Verbeke Mathias**  
KU Leuven  
Campus Brugge



**Dr. Ir. Meert Wannes**  
KU Leuven  
Campus Leuven



**Dr. Feremans Len**  
U Antwerpen

# ANOMALY DETECTION

## severe consequences

Machine breakdown



Financial losses



Production halt



Safety-critical situations

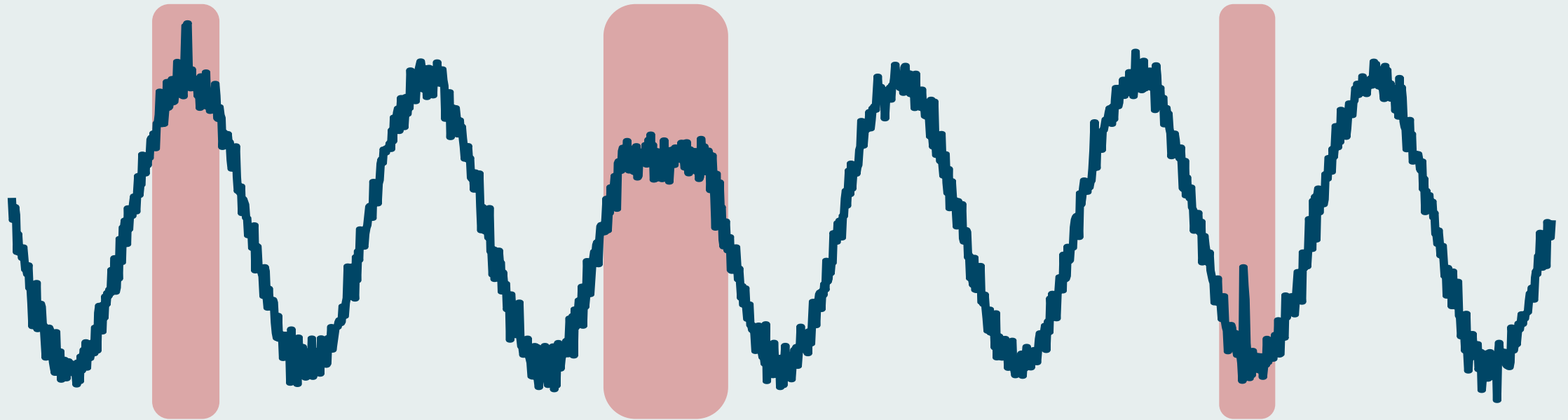


# ANOMALY DETECTION

## in time series data

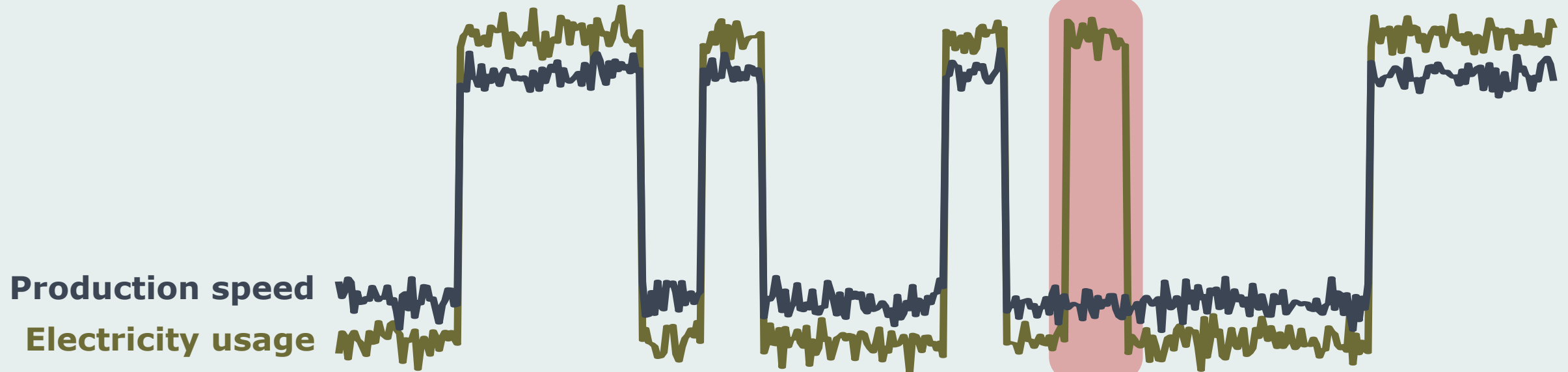
***A TIME SERIES IS AN ORDERED SEQUENCE OF MEASUREMENTS***

***AN ANOMALY IS AN EVENT THAT DEVIATES WHAT IS EXPECTED***



# ANOMALY DETECTION

context is necessary



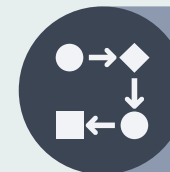
## THE CONTEXT IS A SUBSET OF THE FEATURE SPACE



Operating mode



Time product X is being produced



Manufacturing chain

# ANOMALY DETECTION

## contextual diversity



**INTRA-SPATIAL  
CONTEXT**

**Subset of the values  
within the system**



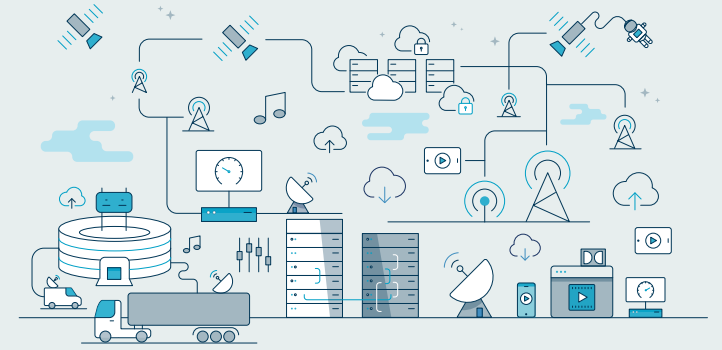
**TEMPORAL  
CONTEXT**

**A subset of the time  
domain (time intervals)**



**INTER-SPATIAL  
CONTEXT**

**Subset of the values within  
other, connected systems**



# CONTEXT-AWARE ANOMALY DETECTION

## CONTEXT IDENTIFICATION

EXPLICITLY DETECT THE RELEVANT CONTEXTUAL INFORMATION



## ANOMALY DETECTION

WRAPPER APPROACH TO INJECT THE RELEVANT CONTEXT

# **TIME SERIES SEMANTIC SEGMENTATION**

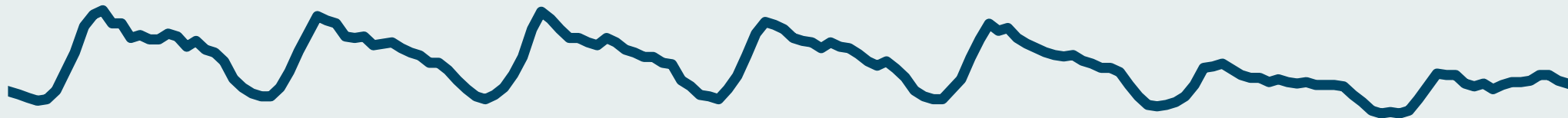
**IDENTIFYING THE TEMPORAL CONTEXT**



# TIME SERIES SEMANTIC SEGMENTATION

***TIME SERIES SEMANTIC SEGMENTATION* AIMS AT IDENTIFYING THE TIME INTERVALS THAT HAVE A SIMILAR MEANING IN THE APPLICATION DOMAIN, IN AN UNSUPERVISED MANNER**

## THE TIME SERIES DATA

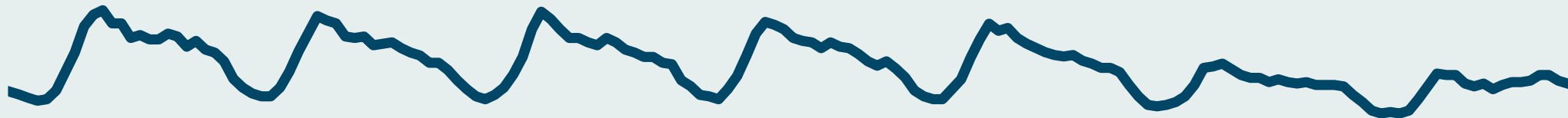


**Worldwide interest in the search term 'mail' in Google over the period of one week, from Monday 20/02/2023 until Sunday 26/02/2023  
(<https://trends.google.com>)**

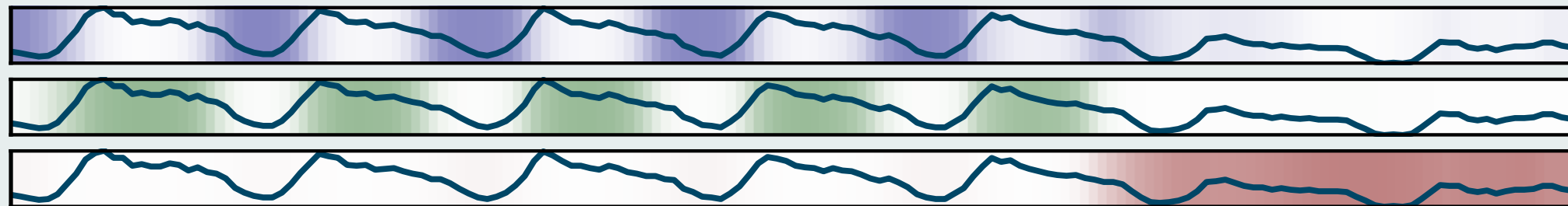
# TIME SERIES SEMANTIC SEGMENTATION

***TIME SERIES SEMANTIC SEGMENTATION* AIMS AT IDENTIFYING THE TIME INTERVALS THAT HAVE A SIMILAR MEANING IN THE APPLICATION DOMAIN, IN AN UNSUPERVISED MANNER**

**THE TIME SERIES DATA**



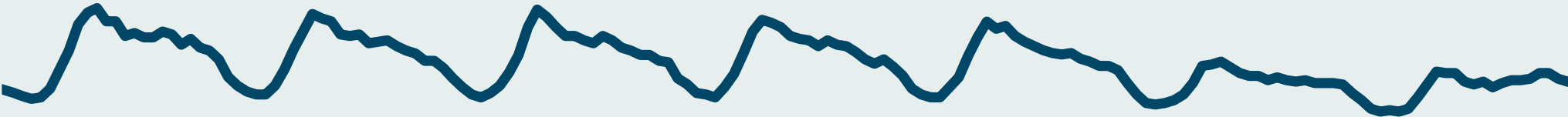
**THE SEMANTIC SEGMENTS ON TOP OF THE TIME SERIES DATA**



# GRADUAL STATE TRANSITIONS

## LEARN A DISTRIBUTION OVER THE SEGMENTS

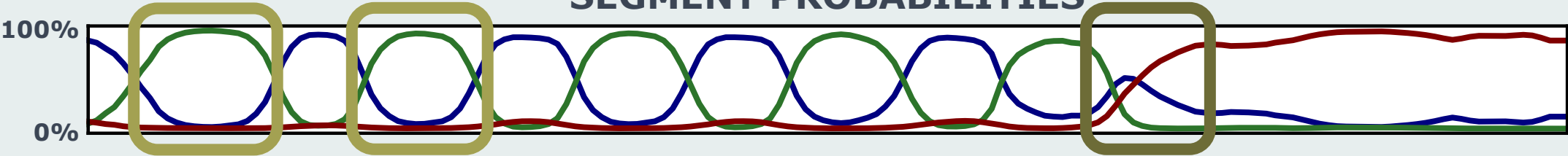
THE TIME SERIES DATA



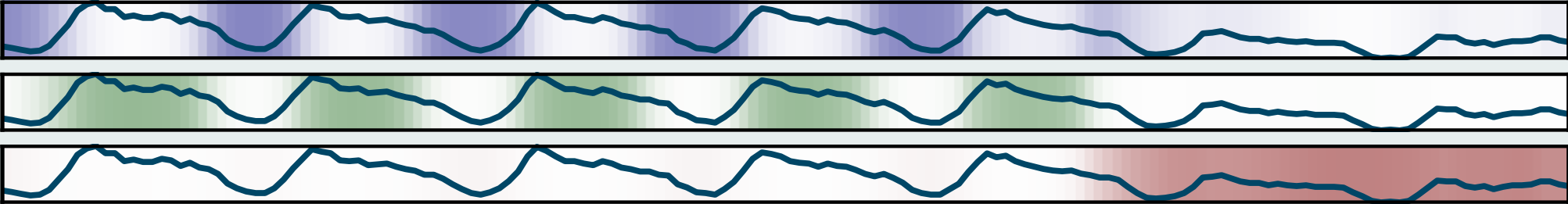
REOCCURRING SEGMENTS

SEGMENT PROBABILITIES

GRADUAL TRANSITIONS

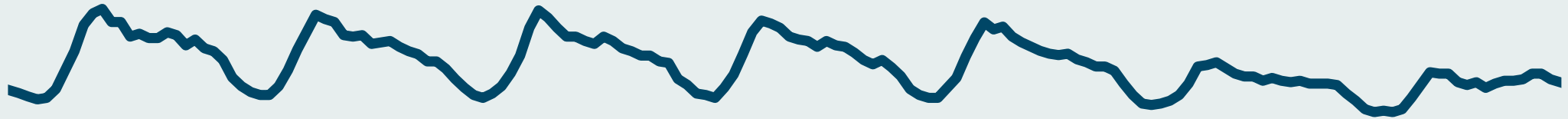


THE SEMANTIC SEGMENTS ON TOP OF THE TIME SERIES DATA

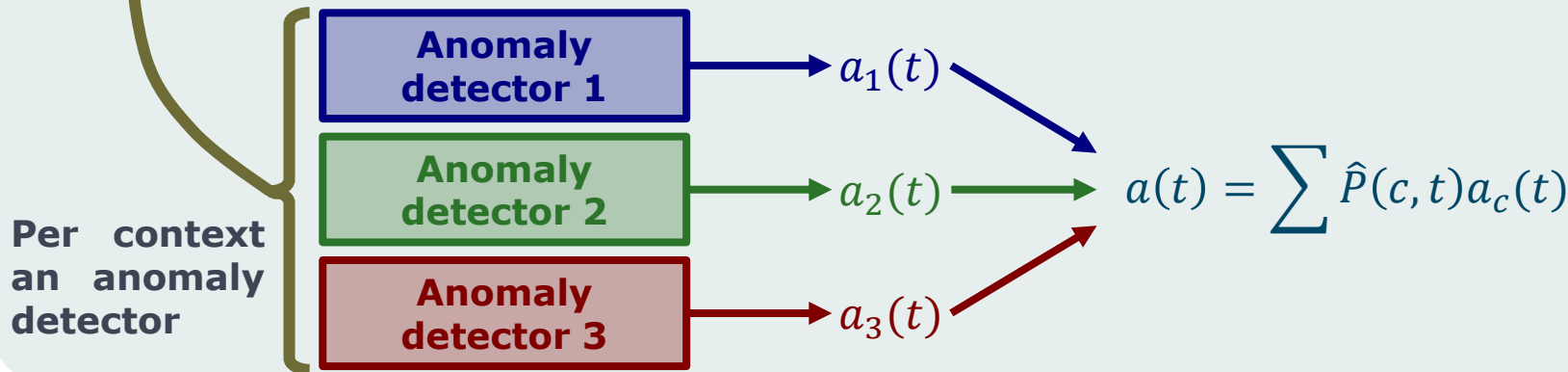
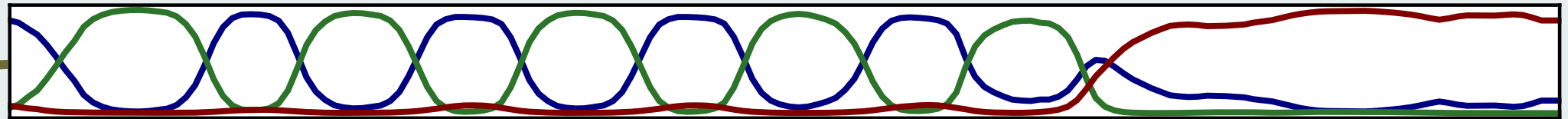


# CONTEXT-AWARE ANOMALY DETECTION

THE TIME SERIES DATA



SEGMENT PROBABILITIES



$a_c(t)$  is the probability of an anomaly occurring at time  $t$  in context  $c$

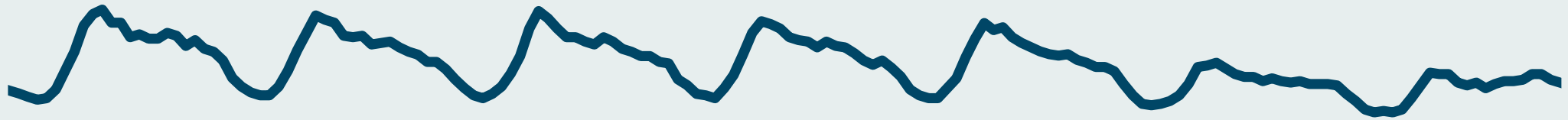
$a(t)$  is the context-aware probability of an anomaly occurring at time  $t$

# **TIME SERIES DISCRETIZATION**

**BECAUSE TIME SERIES DATA IS MESSY**

# MULTI-RESOLUTION WINDOWS

## THE TIME SERIES DATA



## SEGMENTS OF LENGTH 1 DAY



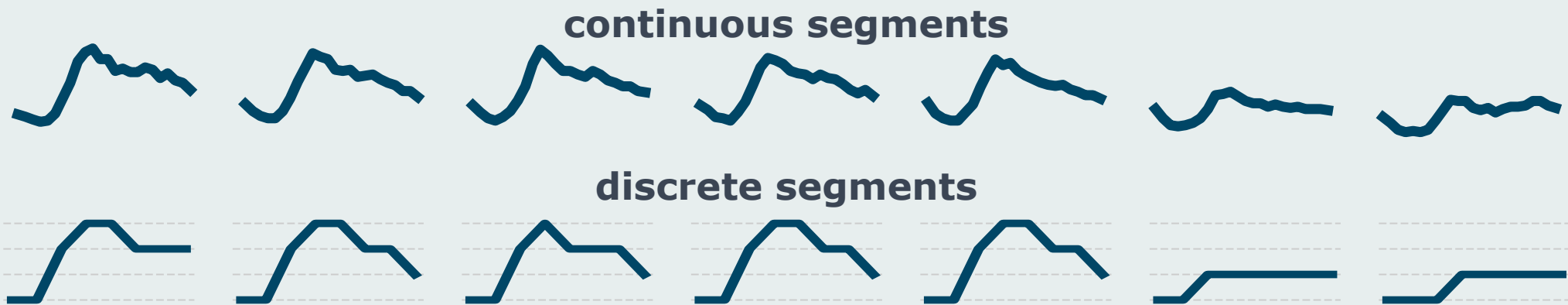
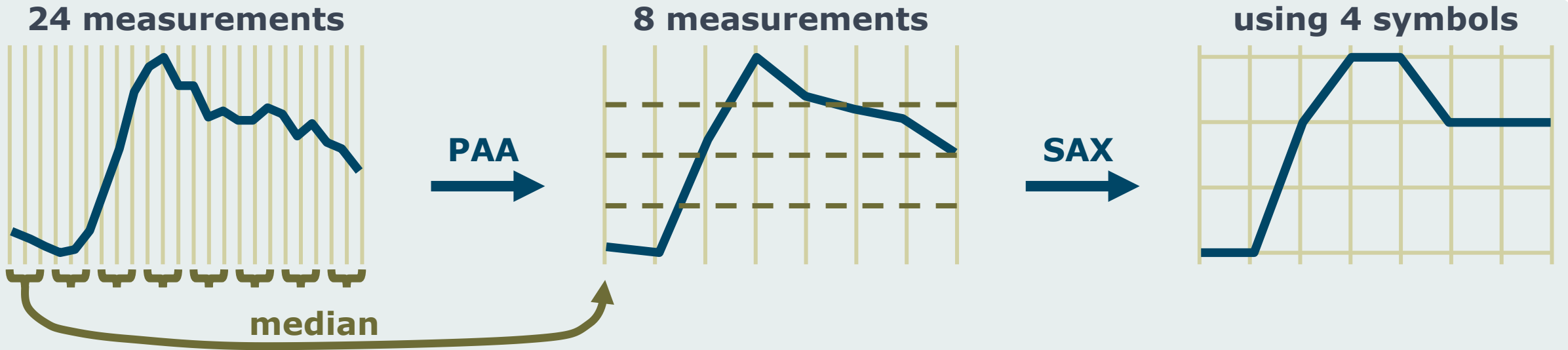
## SEGMENTS OF LENGTH 2 DAYS



Multiple resolutions

Overlap

# DISCRETIZATION



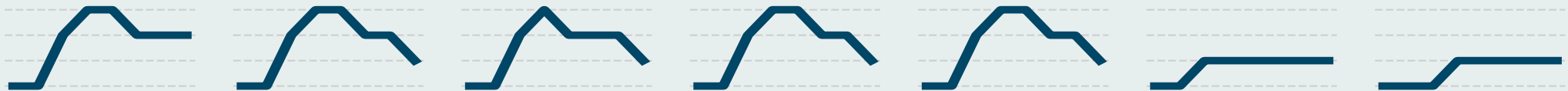
# PaTSS: PATTERN-BASED TIME SERIES SEGMENTATION





# (1) SEQUENTIAL PATTERNS

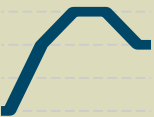

A *SEQUENTIAL PATTERN* IS AN ORDERED SEQUENCE OF DISCRETE SYMBOLS. WE SAY THAT THE PATTERN IS *FREQUENT* IF IT OCCURS IN MANY DISCRETE SUBSEQUENCES

discrete segments



	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.	Sun.	sup <sub>rel</sub>
	X	X	X	X	X			5/7
						X	X	2/7

## (2) PATTERN-BASED EMBEDDING

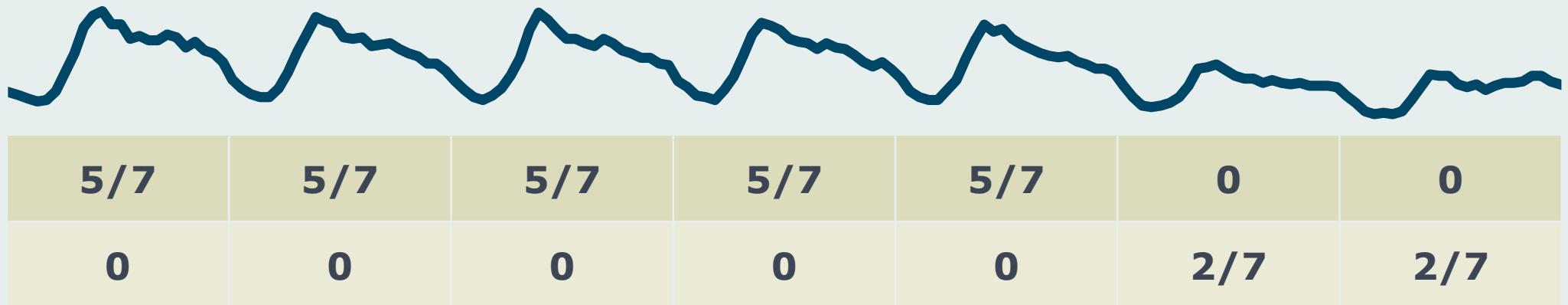
	Mon.	Tue.	Wed.	Thu.	Fri.	Sat.	Sun.	sup <sub>rel</sub>
	X	X	X	X	X			5/7
						X	X	2/7

$$E_t(P) = \frac{|cover(P) \cap cover(t)|}{|cover(t)|} \cdot support_{relative}(P)$$

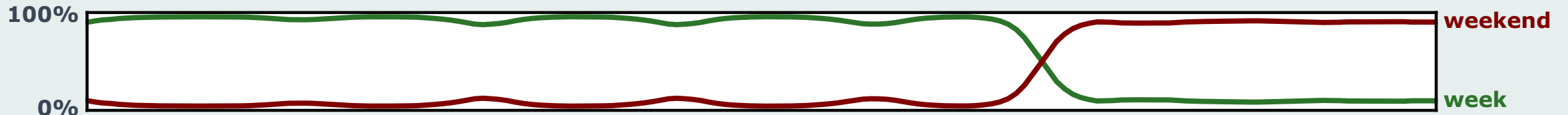
5/7	5/7	5/7	5/7	5/7	5/7	0	0
0	0	0	0	0	0	2/7	2/7

# (3) SEMANTIC SEGMENTATION

THE TIME SERIES DATA



SEGMENT PROBABILITIES



**Logistic regression with supervision from k-means clustering**  
to learn the probability of a certain semantic segment at each time point

# CHICKEN DANCE

## FOUR DANCE MOVES REPEATED TWICE

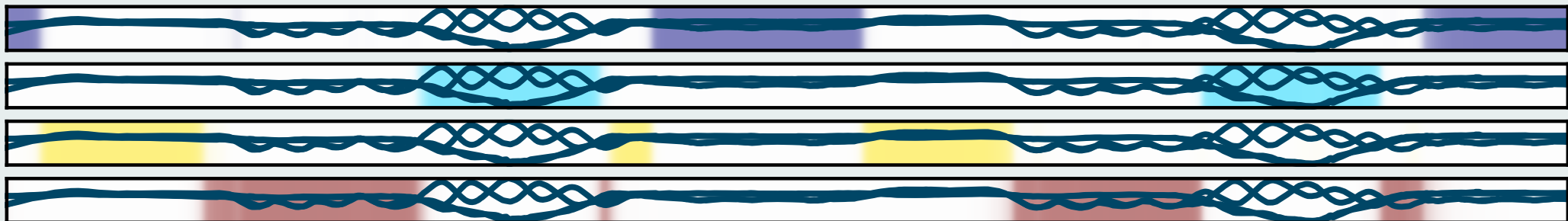
THE TIME SERIES DATA



SEGMENT PROBABILITIES

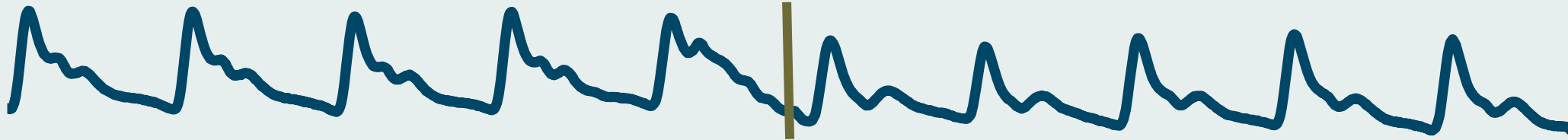


THE SEMANTIC SEGMENTS ON TOP OF THE TIME SERIES DATA

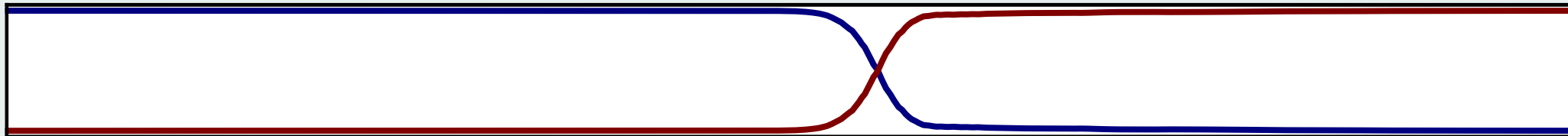


# ARTERIAL BLOOD PRESSURE OF A PERSON LAYING ON A TILT TABLE

THE TIME SERIES DATA



SEGMENT PROBABILITIES



THE SEMANTIC SEGMENTS ON TOP OF THE TIME SERIES DATA

