Challenges in Explainable Artificial Intelligence for Industry 4.0

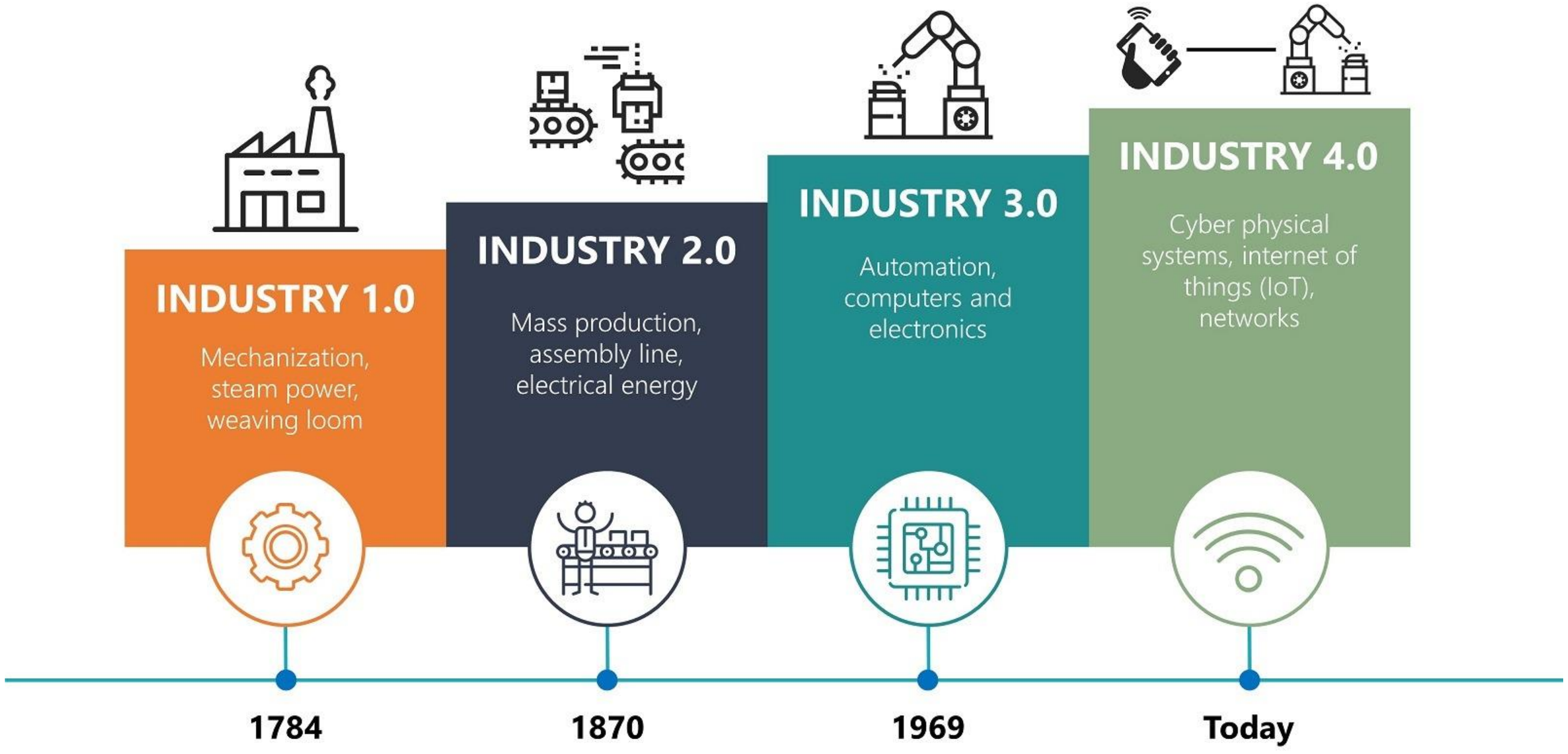
Szymon Bobek

Jagiellonian University
2023

JAGIELLONIAN UNIVERSITY
IN KRAKÓW

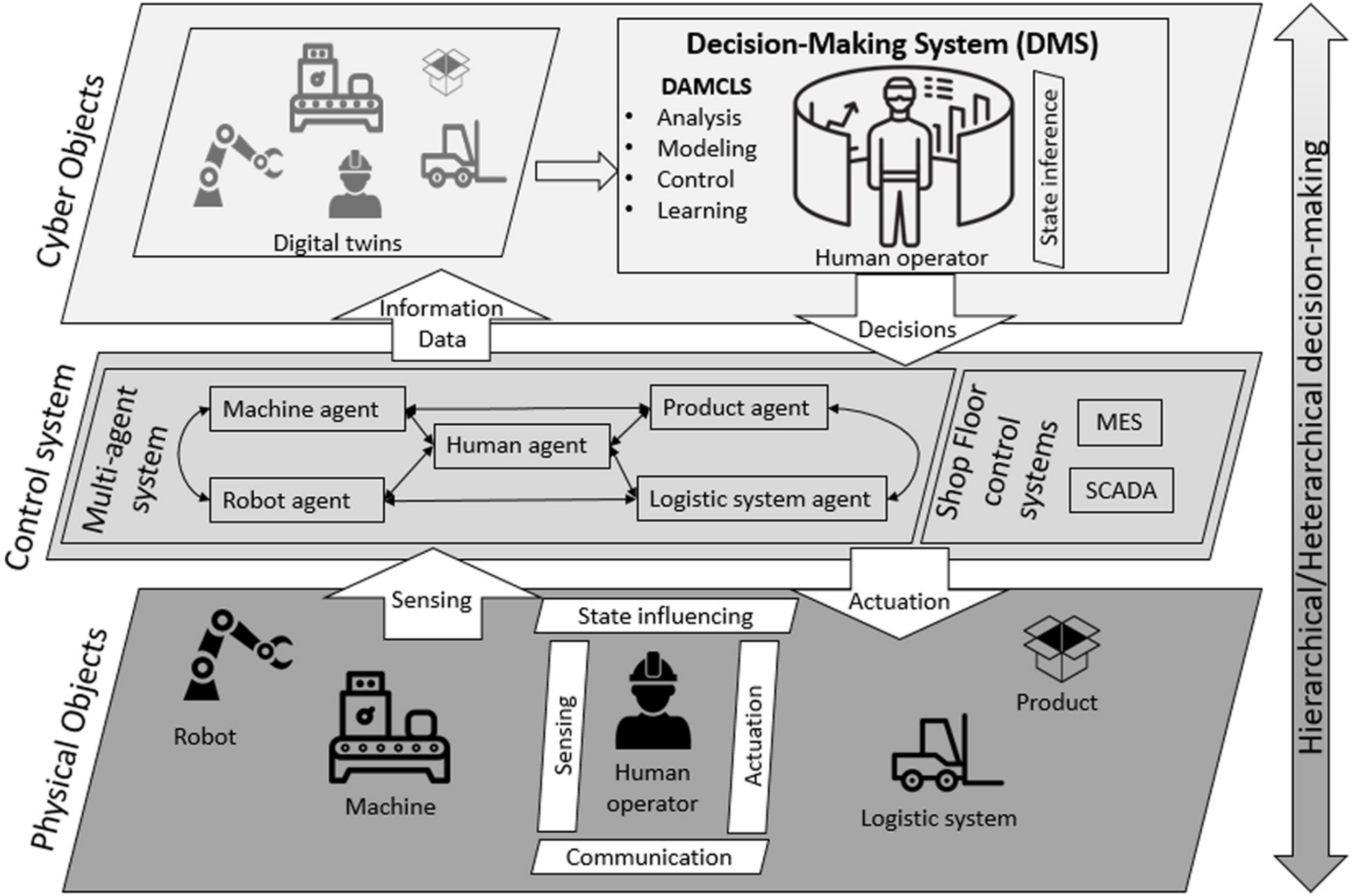
https://geist.re

Industry 4.0

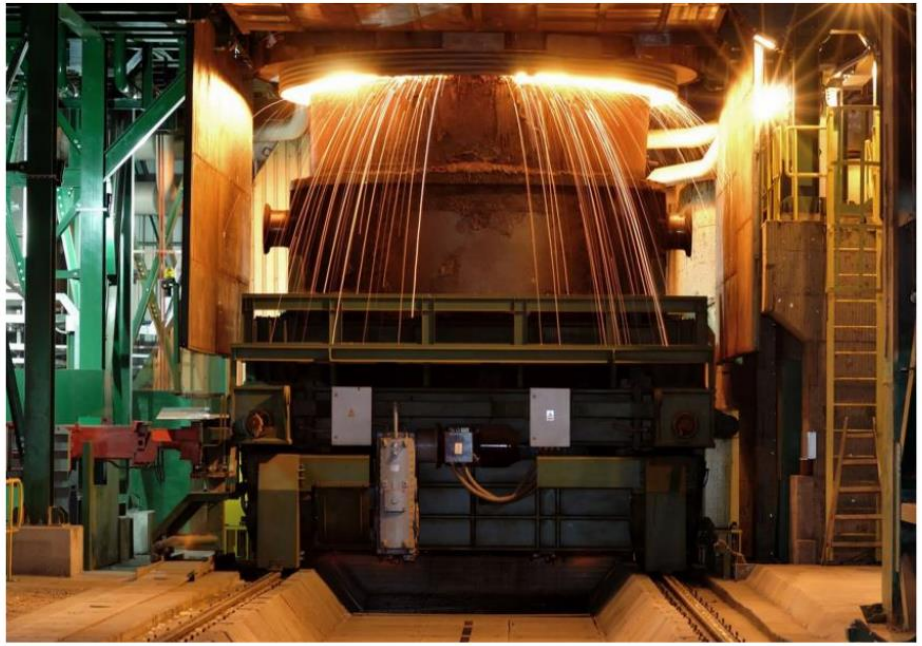


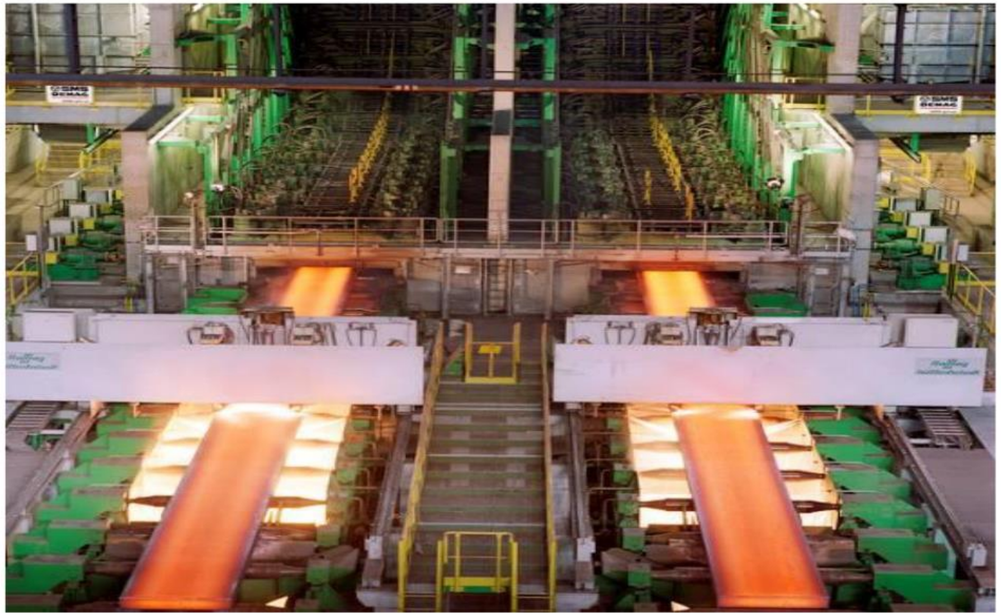
Source: https://www.presentationpoint.com/blog/data-signals-triggers-industry-4-0/ (Accessed 12.01.2021)

Let us have an Industry 4.0 factory!



Source: Cimini, C.; Pirola, F.; Pinto, R.; Cavalieri, S. A human-in-the-loop manufacturing control architecture for the next generation of production systems. Journal of Manufacturing Systems 2020, 54, 258–271. doi:https://doi.org/10.1016/j.jmsy.2020.01.002.





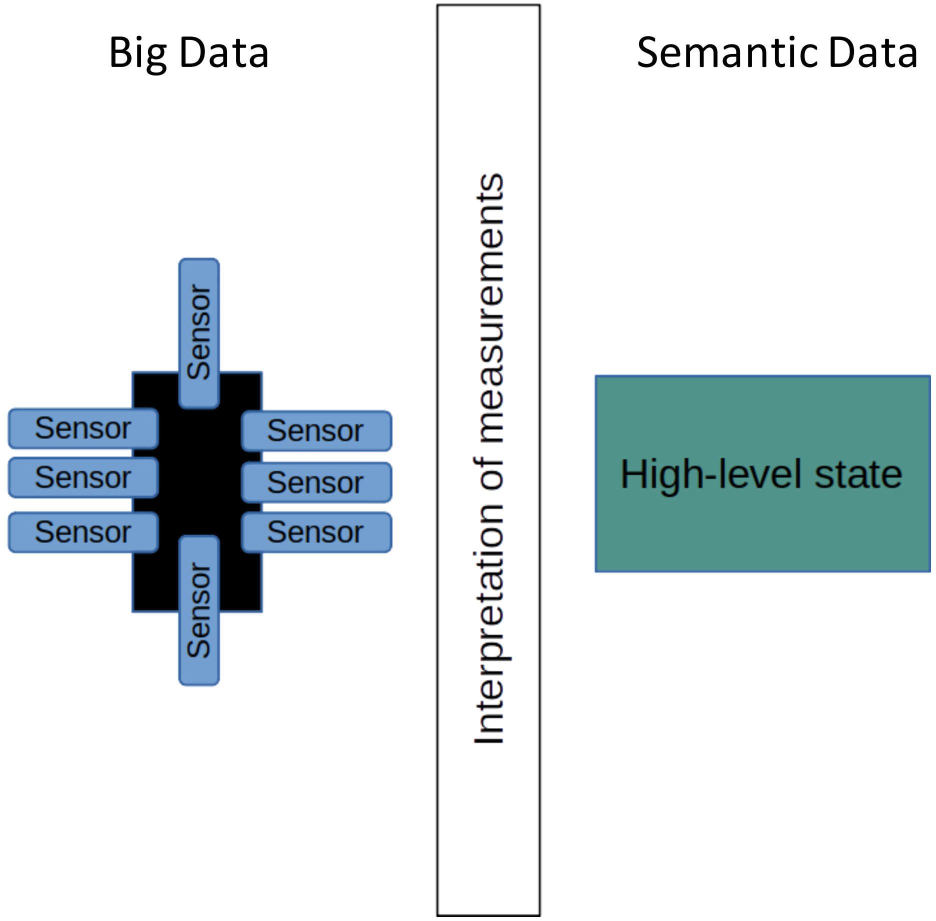
Source: Presentation @ DSAA2021


Source: https://www.mdpi.com/1996-1073/13/24/6630/htm

Working with experts

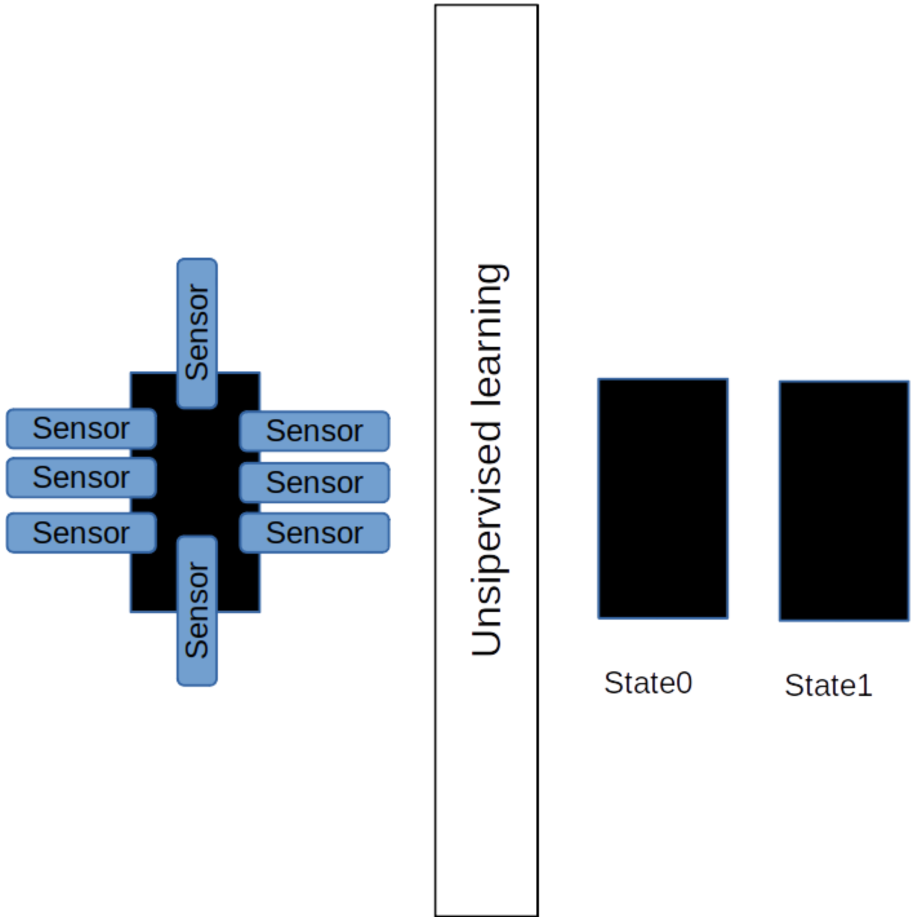
Knowledge discovery from industrial logs

- PACMEL Project (http://pacmel.geist.re)
- Industry 4.0: everything is measured
- Low-level measurements to higher-level states
- Expert-defined states vs. Automatically discovered states
  - Data comes with no labels
  - Expert may be to general, or too specific
  - There is a lot of measurements



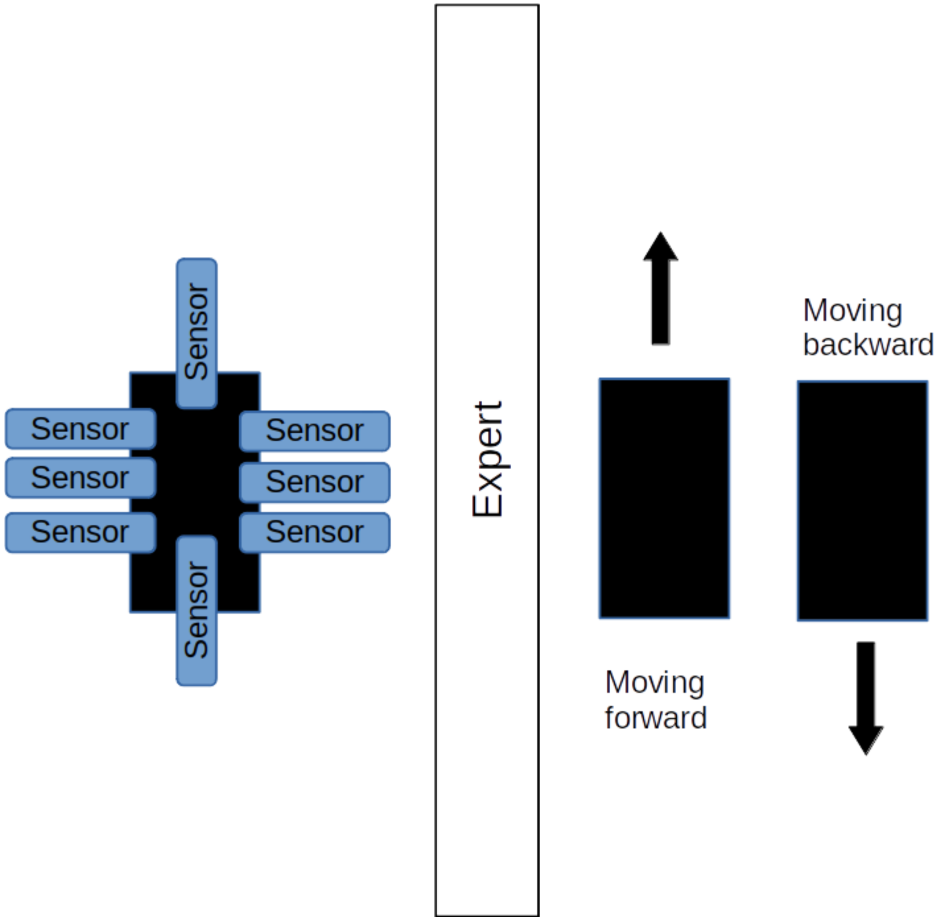
Knowledge discovery from industrial logs

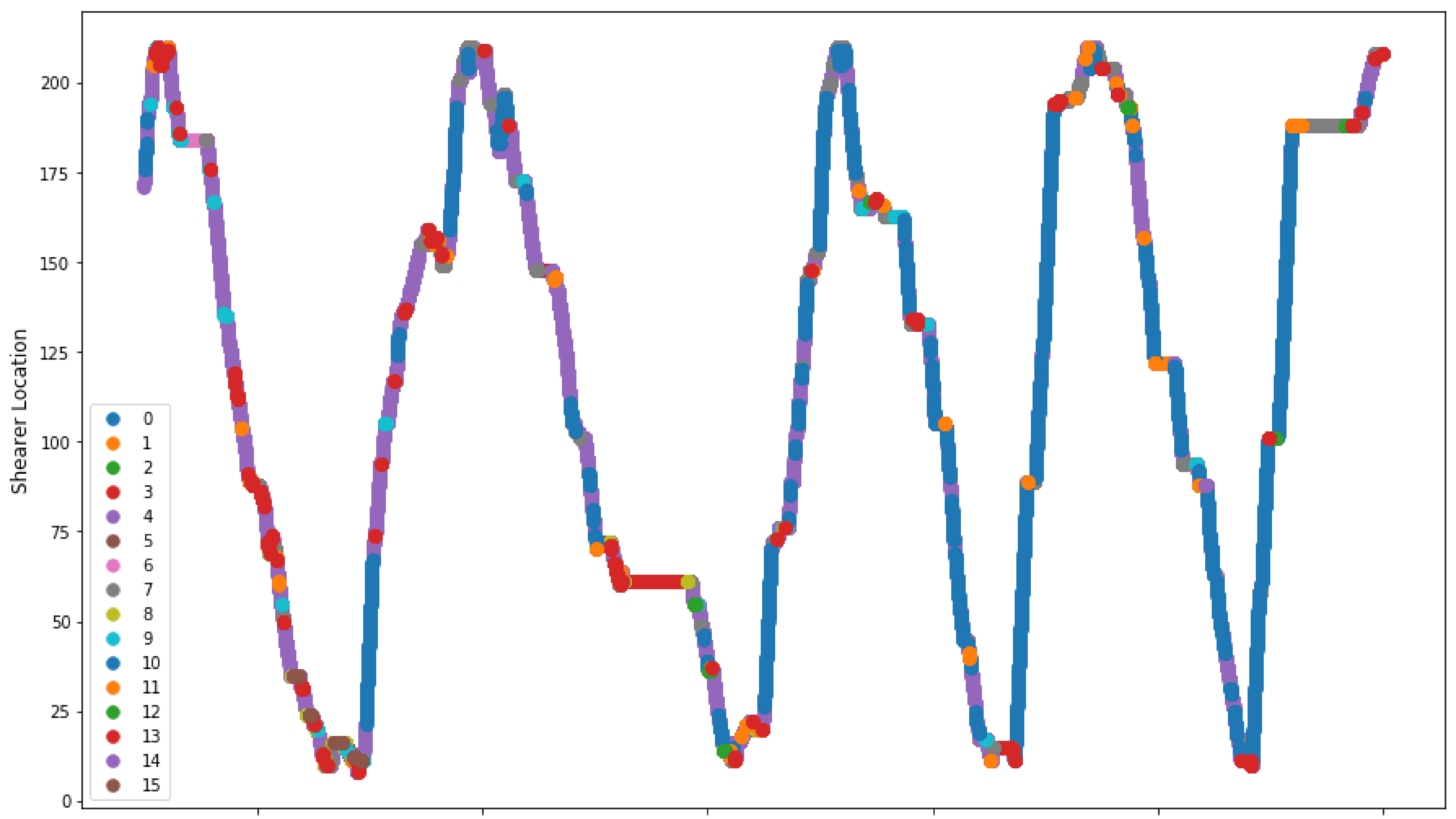
- PACMEL Project (http://pacmel.geist.re)
- Industry 4.0: everything is measured
- Low-level measurements to higher-level states
- Expert-defined states vs. Automatically discovered states
  - Data comes with no labels
  - Expert may be to general, or too specific
  - There is a lot of measurements



Knowledge discovery from industrial logs

- PACMEL Project (http://pacmel.geist.re)
- Industry 4.0: everything is measured
- Low-level measurements to higher-level states
- Expert-defined states vs. Automatically discovered states
  - Data comes with no labels
  - Expert may be to general, or too specific
  - There is a lot of measurements





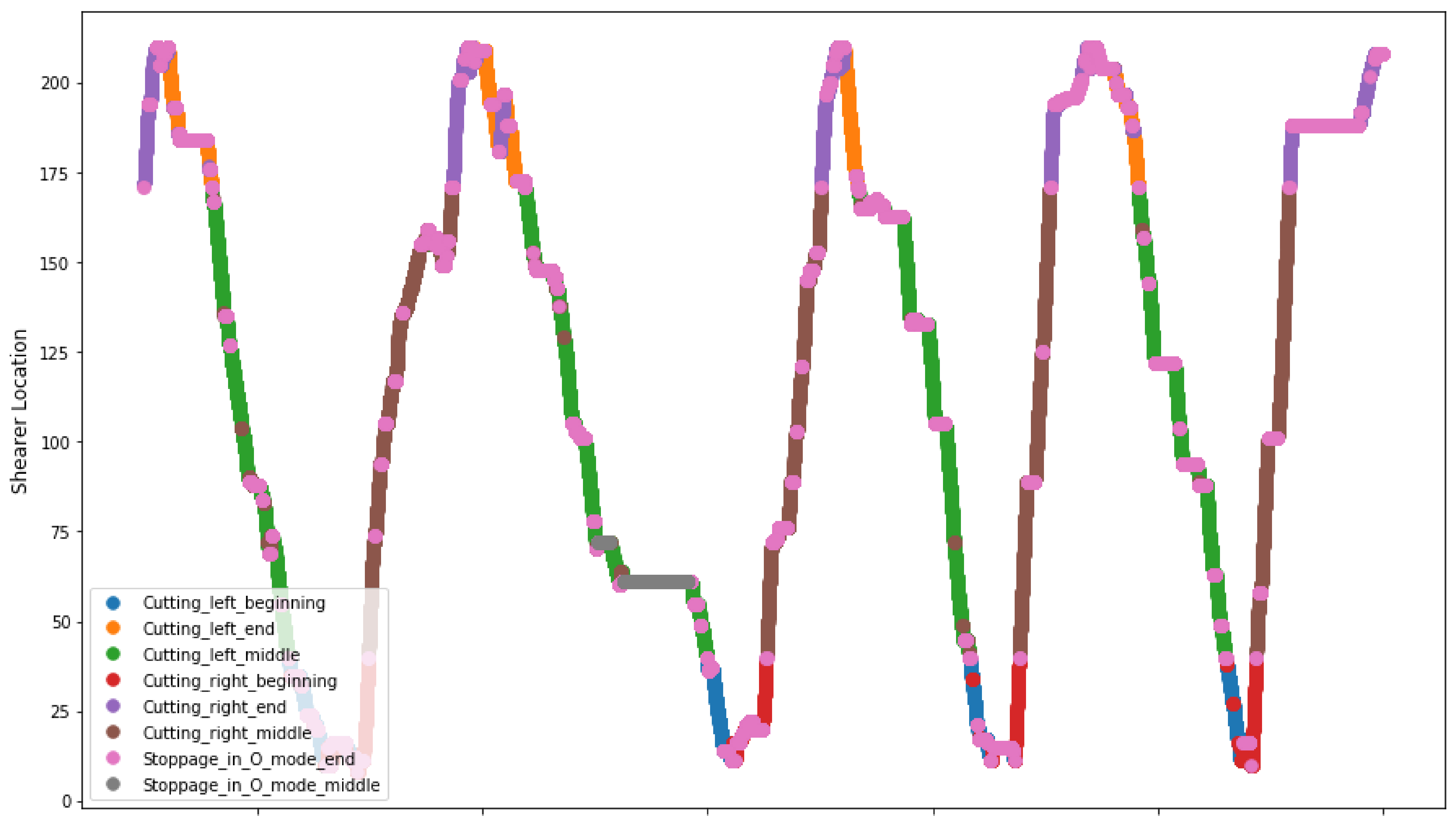
Shearer Location
0
25
50
75
100
125
150
175
200
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

Theoretical states are given by the expert

					State	#
= 0	> 0	= 0	= 1	= any	set movingLeft	1
= 0	> 0	= 0	= 0	= any	set movingRight	1
> 0	> 0	= any	= 1	< 40	set cuttingLeftBegining	1
> 0	> 0	= any	= 1	∈ [40 .. 180]	set cuttingLeftMiddle	1
> 0	> 0	= any	= 1	>= 180	set cuttingLeftEnd	1
> 0	> 0	= any	= 0	< 40	set cuttingRightBeginning	1
> 0	> 0	= any	= 0	∈ [40 .. 180]	set cuttingRightMiddle	1
> 0	> 0	= any	= 0	> 180	set cuttingRightEnd	1
= any	= any	= any	≠ any	< 40	set stoppageInOModeBeginn…	1
= any	= any	= any	≠ any	< 180	set stoppageInOModeMiddle	1
= any	= any	= any	≠ any	>= 180	set stoppageInOModeEnd	1

expertLabeling

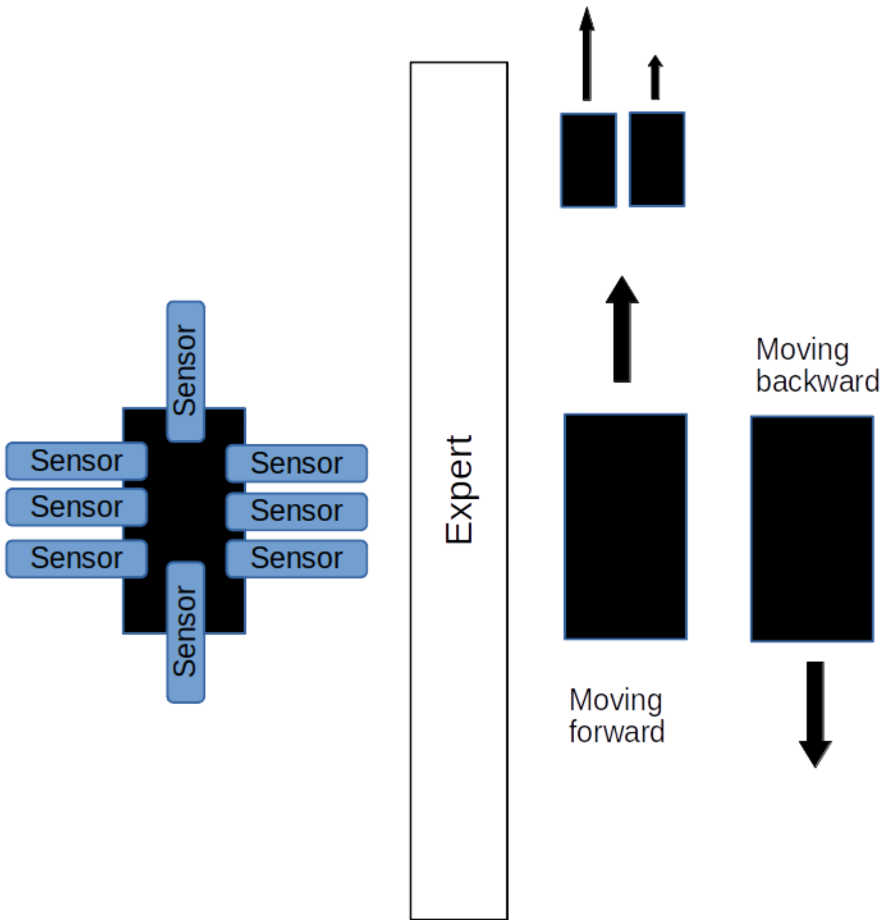
Add condition Add decision Add rule



Shearer Location
200
175
150
125
100
75
50
25
0
Cutting_left_beginning
Cutting_left_end
Cutting_left_middle
Cutting_right_beginning
Cutting_right_end
Cutting_right_middle
Stoppage_in_O_mode_end
Stoppage_in_O_mode_middle

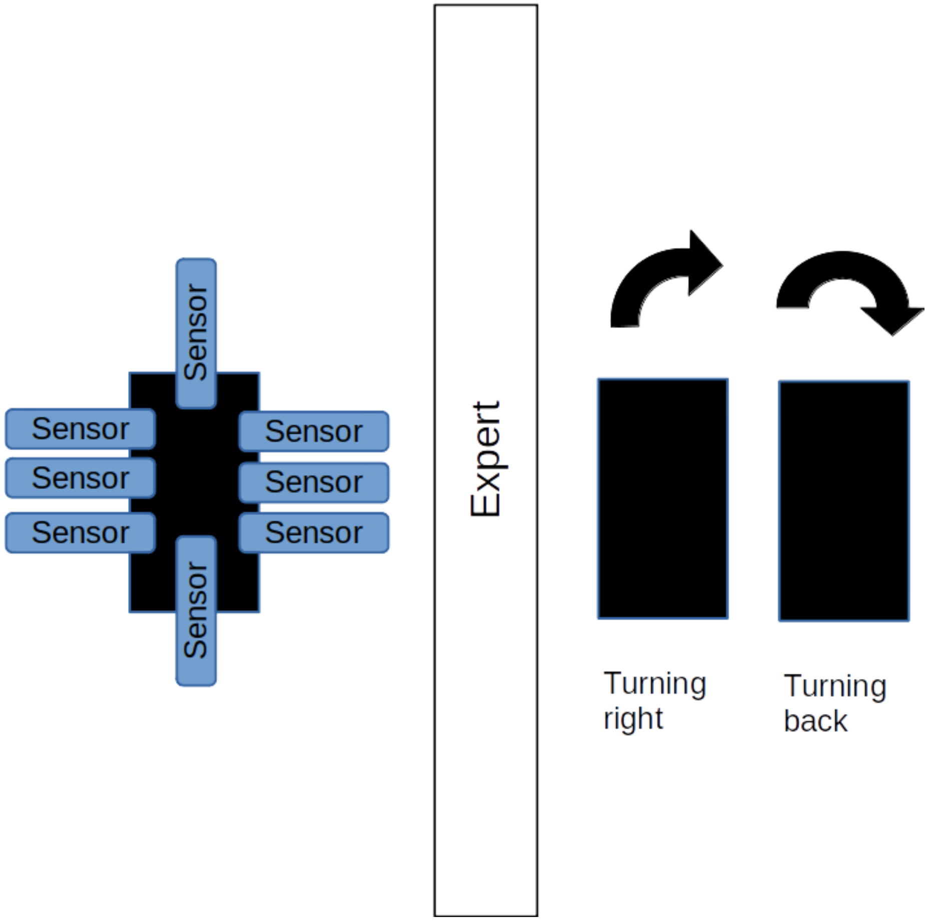
Knowledge discovery from industrial logs

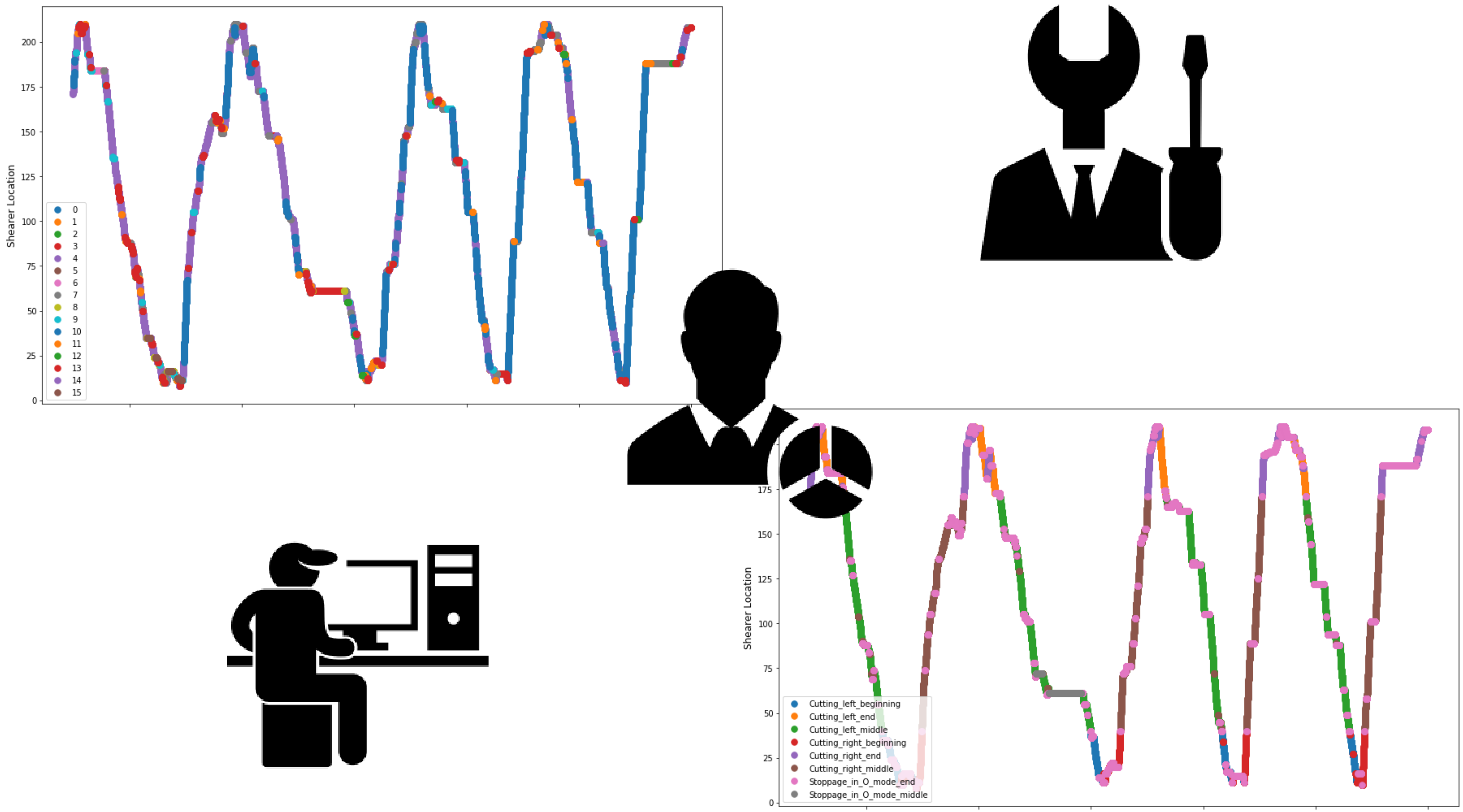
- PACMEL Project (http://pacmel.geist.re)
- Industry 4.0: everything is measured
- Low-level measurements to higher-level states
- Expert-defined states vs. Automatically discovered states
  - Data comes with no labels
  - Expert may be to general, or too specific
  - There is a lot of measurements



Knowledge discovery from industrial logs

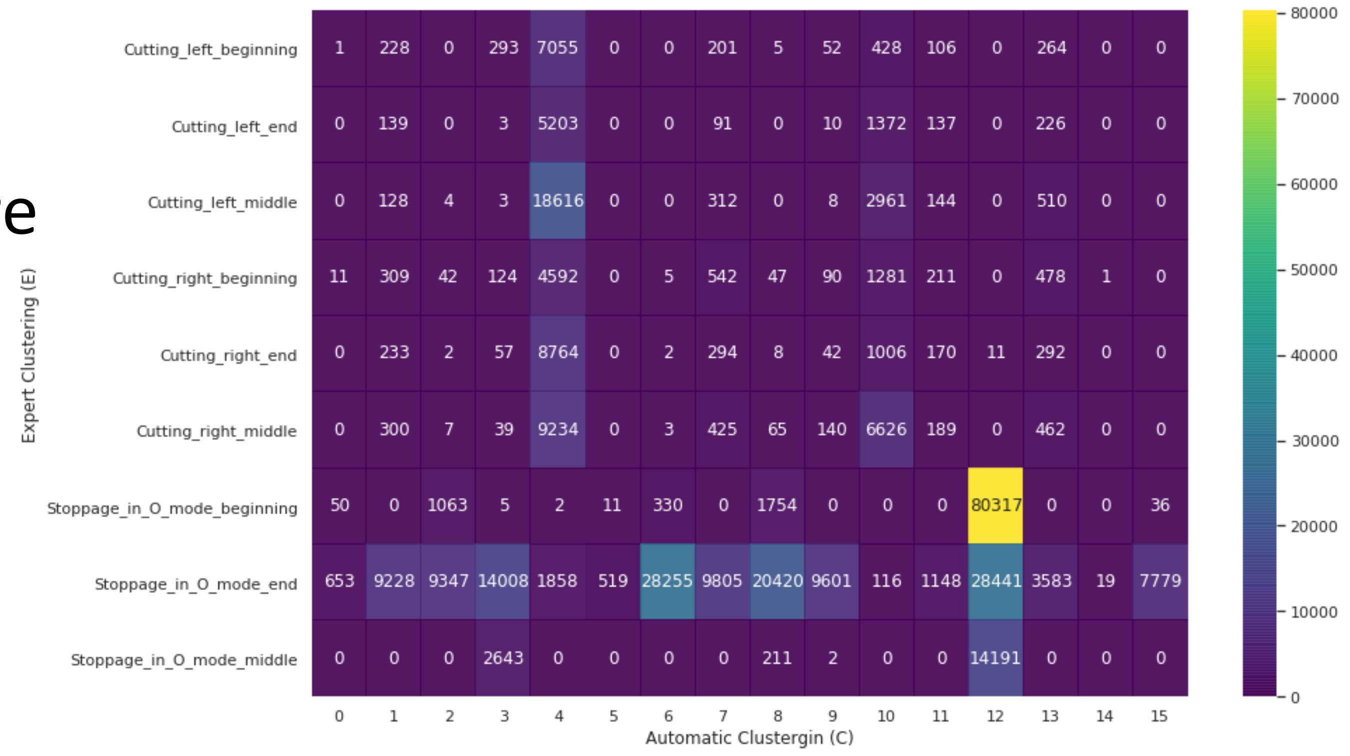
- PACMEL Project (http://pacmel.geist.re)
- Industry 4.0: everything is measured
- Low-level measurements to higher-level states
- Expert-defined states vs. Automatically discovered states
    - Data comes with no labels
    - Expert may be to general, or too specific
    - There is a lot of measurements



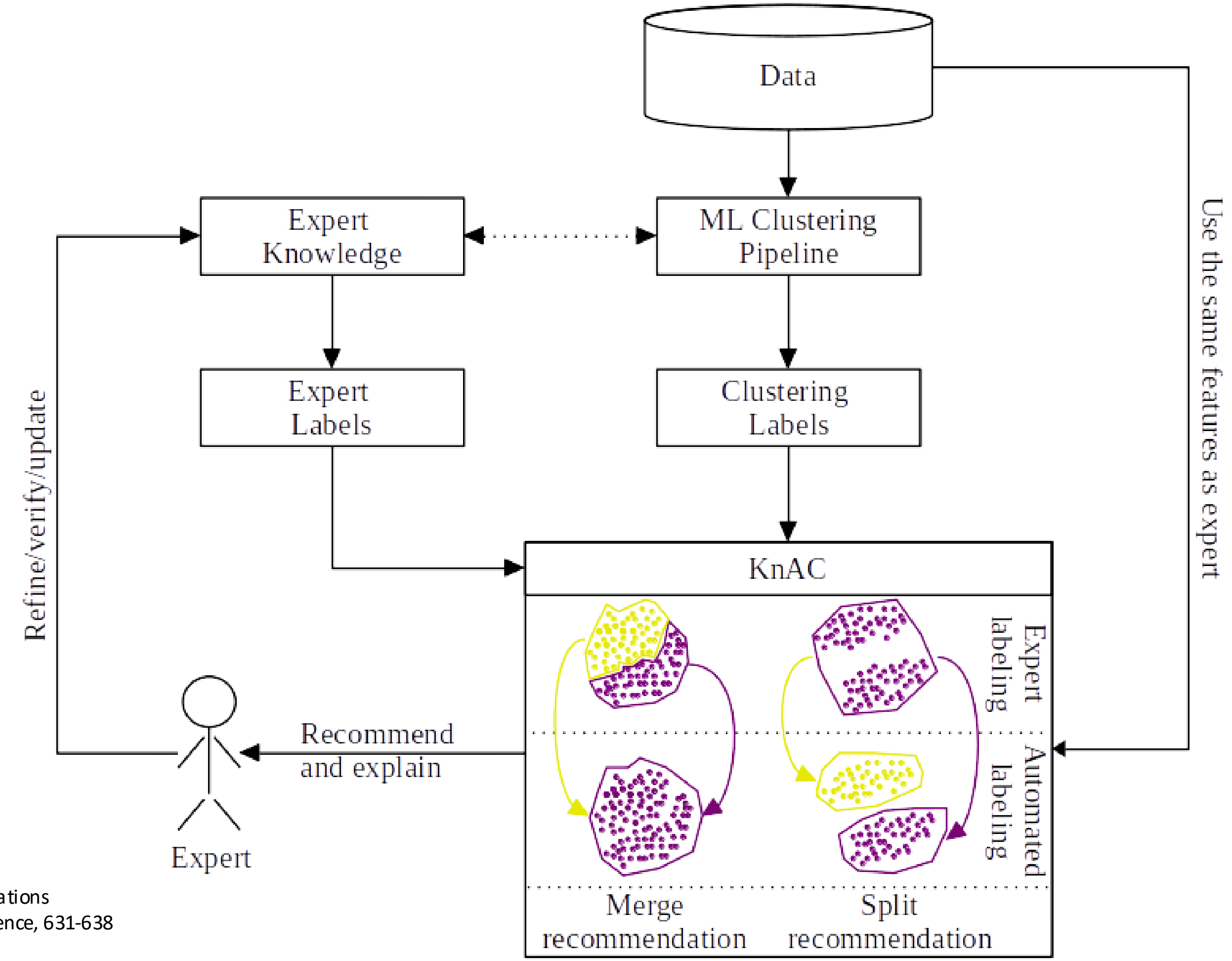

Shearer Location
200
175
150
125
100
75
50
25
0
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Shearer Location
175
150
125
100
75
50
25
0
Cutting_left_beginning
Cutting_left_end
Cutting_left_middle
Cutting_right_beginning
Cutting_right_end
Cutting_right_middle
Stoppage_in_O_mode_end
Stoppage_in_O_mode_middle

How to confront expert and automatic labelling?

- Analysis of each of the states separately via contingency matrix
- Adjusted rand score
- Adjusted mutual info score
- Homogeneity
- Consistency
- V-measure
- …



Knowledge Augmented Clustering (KnAC)

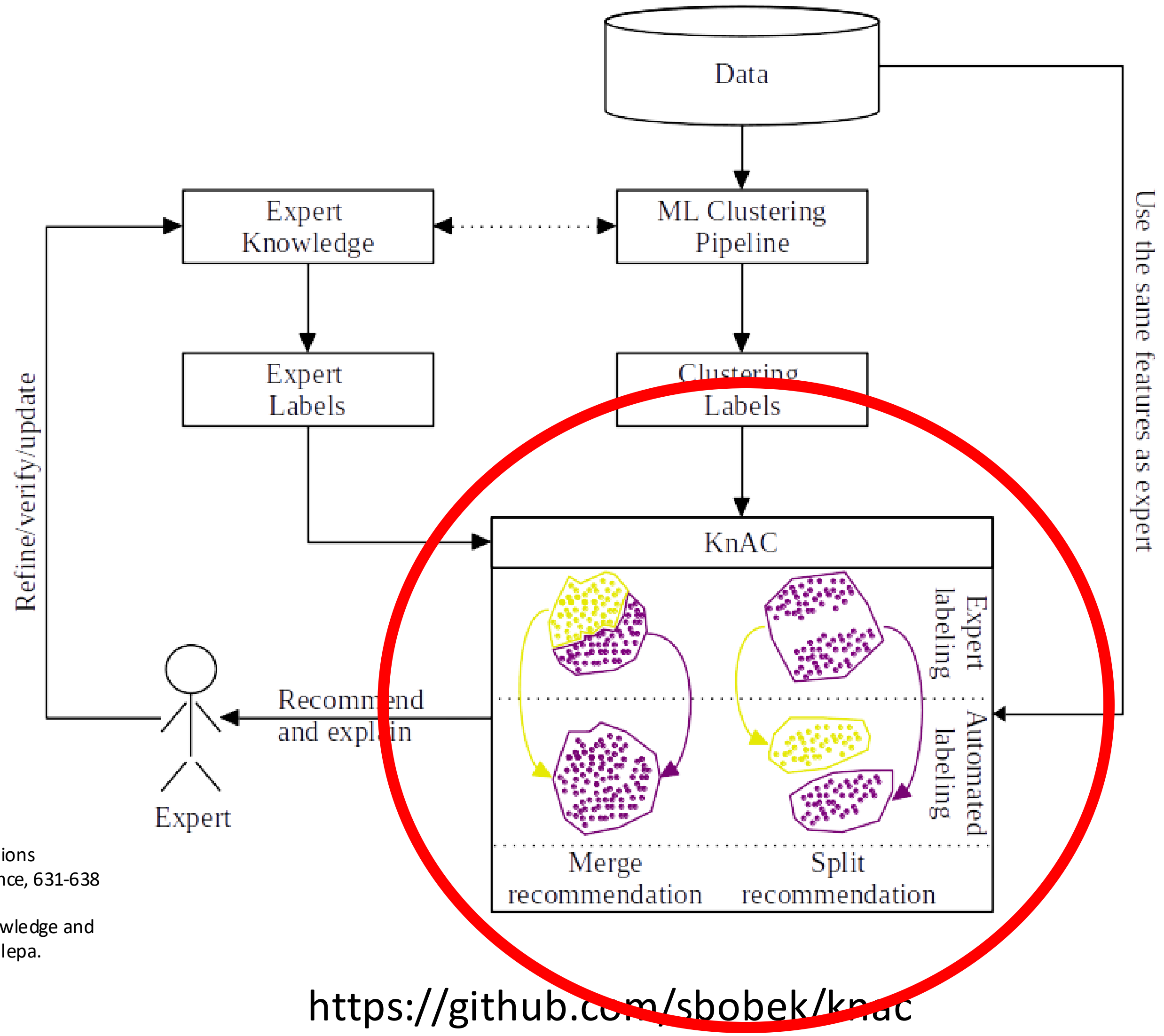


Augmenting Au tomatic Clustering with Expert Knowledge and Explanations
S. Bobek, G.J. Nalepa, International Conference on Computational Science, 631-638

KnAC: an approach for enhancing cluster analysis with background knowledge and explanations, S. Bobek, M. Kuk, J. Brzegowski, E. Brzychczy, and G. J. Nalepa.
ArXiv: https://arxiv.org/abs/2112.08759

https://github.com/sbobek/knac

Knowledge Augmented Clustering (KnAC)

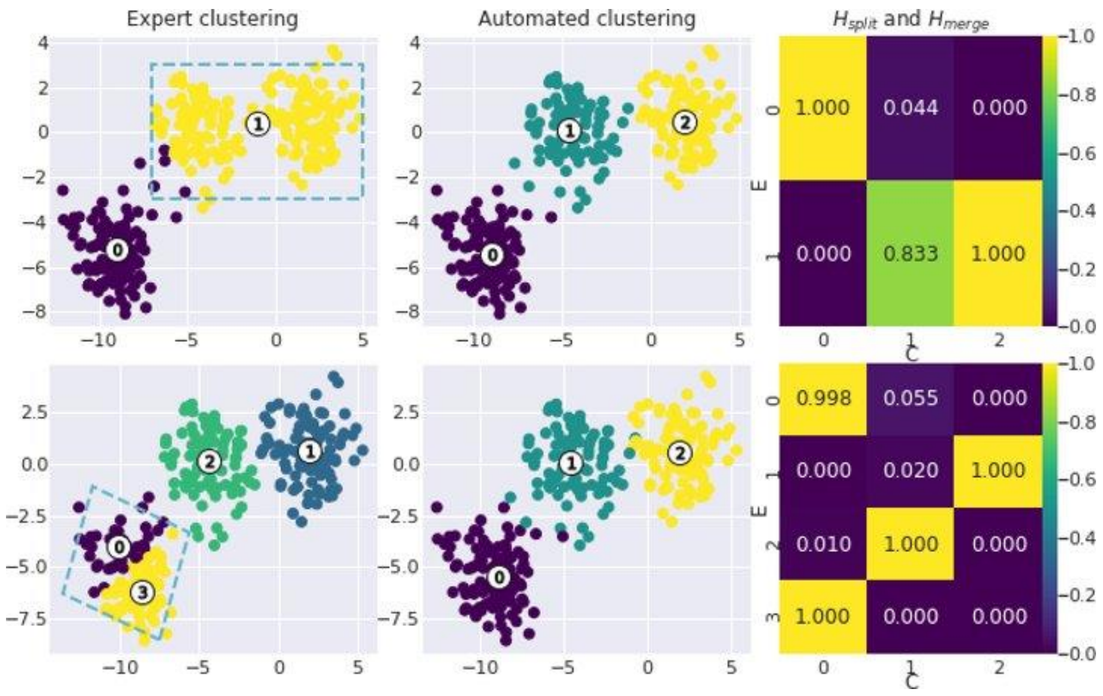


Augmenting Automatic Clustering with Expert Knowledge and Explanations
S. Bobek, G.J. Nalepa, International Conference on Computational Science, 631-638

KnAC: an approach for enhancing cluster analysis with background knowledge and explanations, S. Bobek, M. Kuk, J. Brzegowski, E. Brzychczy, and G. J. Nalepa.
ArXiv: https://arxiv.org/abs/2112.08759

https://github.com/sbobek/knac

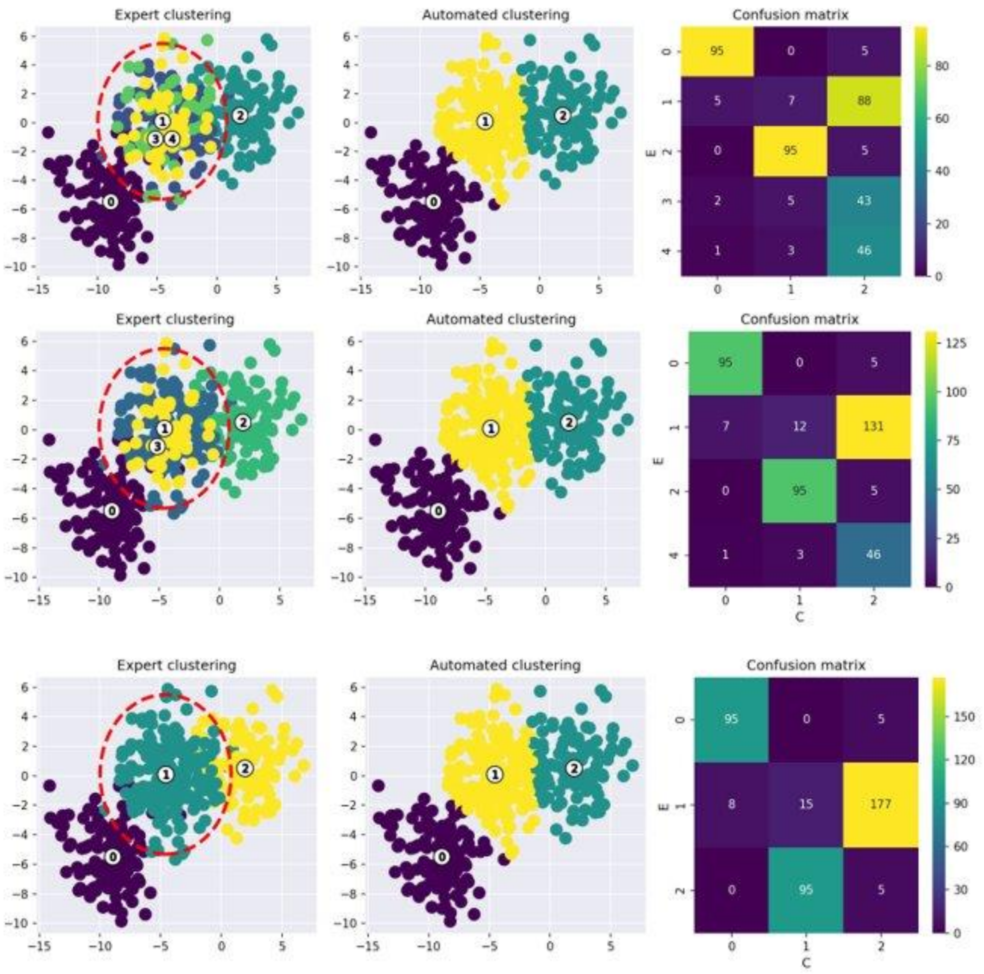
Expert labels vs. Clustering labels



- Split expert clusters into more specific ones
- Merge expert clusters that seem to be similar
- It is an iterative approach

https://github.com/sbobek/knac

Expert labels vs. Clustering labels



- Split expert clusters into more specific ones
- Merge expert clusters that seem to be simillar
- It is an iterative approach

https://github.com/sbobek/knac

Splitting expert cluster

- We calculated entropy of each cluster distribution with respect to expert labels
- We scaled rows of distribution matrix to deal with different sized expert clusters
- We divided normalized matrix with entropy values
- The split confidence was calculated by averaging each row of such matrix

	C1	C2	C3	C4
Expert Cluster 1	400	0	1	..
Expert Cluster 2	1000	1000	1000	..
Expert Cluster 3	200	0	3	..
Expert Cluster 4	600	0	10	..

$$H_{i,j}^{split} = \frac{M_{i,j}}{||M_i||_2 \left[\frac{H(M_i)}{log2(||E||)} + 1 \right]}$$

Splitting expert cluster

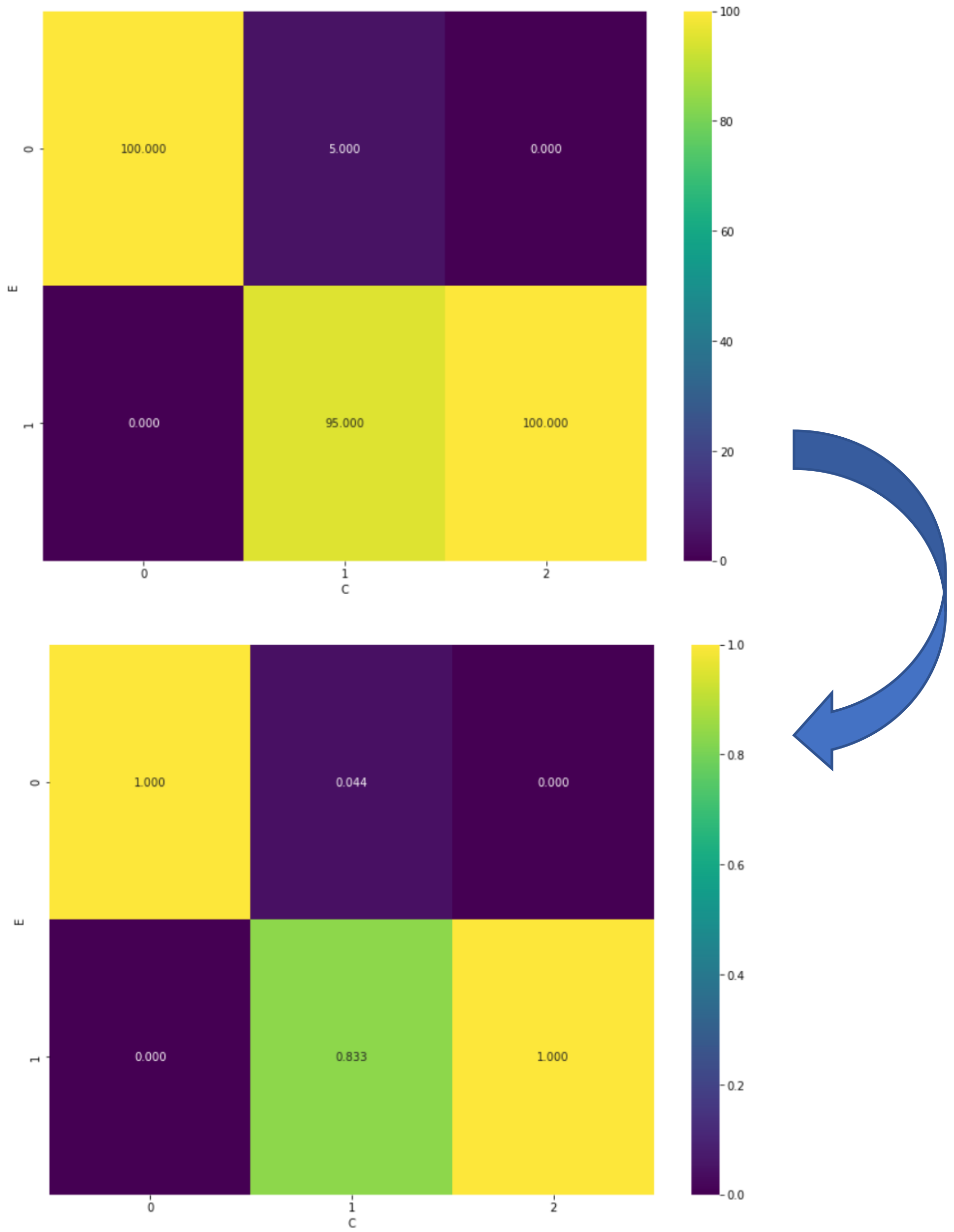
- We calculated entropy of each cluster distribution with respect to expert labels
- We scaled rows of distribution matrix to deal with different sized expert clusters
- We divided normalized matrix with entropy values
- The split confidence was calculated by averaging each row of such matrix

	C1	C2	C3	C4
Expert Cluster 1	400	0	1	..
Expert Cluster 2	1000	0	1000	0
Expert Cluster 3	0	0	3	..
Expert Cluster 4	0	60	0	60

$$H_{i,j}^{split} = \frac{M_{i,j}}{||M_i||_2 \left[\frac{H(M_i)}{log2(||E||)} + 1 \right]}$$

Splitting expert cluster

- We calculated entropy of each cluster distribution with respect to expert labels
- We scaled rows of distribution matrix to deal with different sized expert clusters
- We divided the scaled matrix with entropy values
- The split confidence was calculated by averaging each row of such matrix



Merging expert cluster

- We calculated *l2* normalized distribution matrix
- We calculated cosine similarity between rows to denote expert clusters that were similarly splitted with automated method

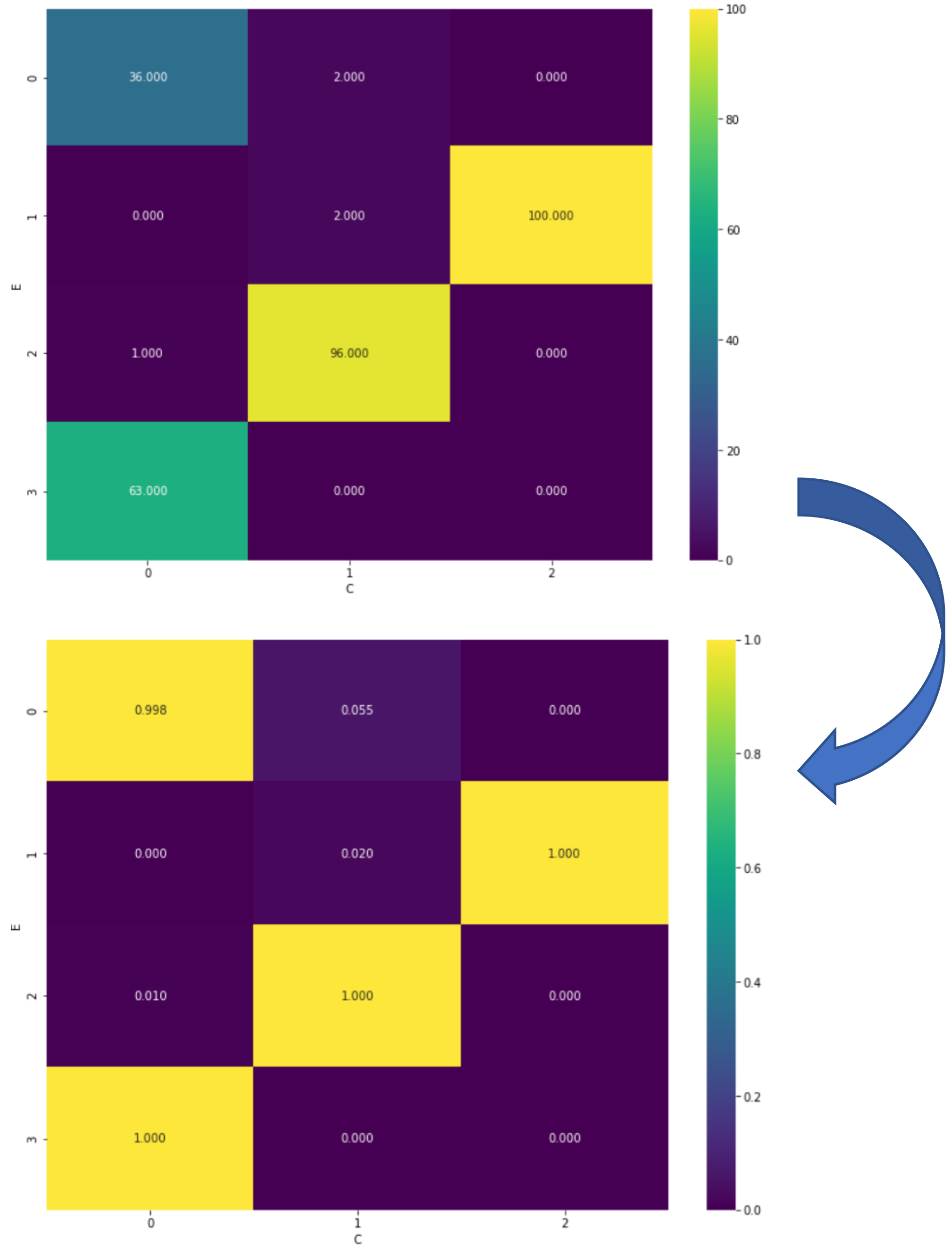
	C1	C2	C3	C4
Expert Cluster 1	400	0	1	..
Expert Cluster 2	1000	0	1000	0
Expert Cluster 3	0	0	3	..
Expert Cluster 4	60	0	60	0

$$H_{i,j}^{merge} = \frac{M_{i,j}}{||M_i||_2}$$

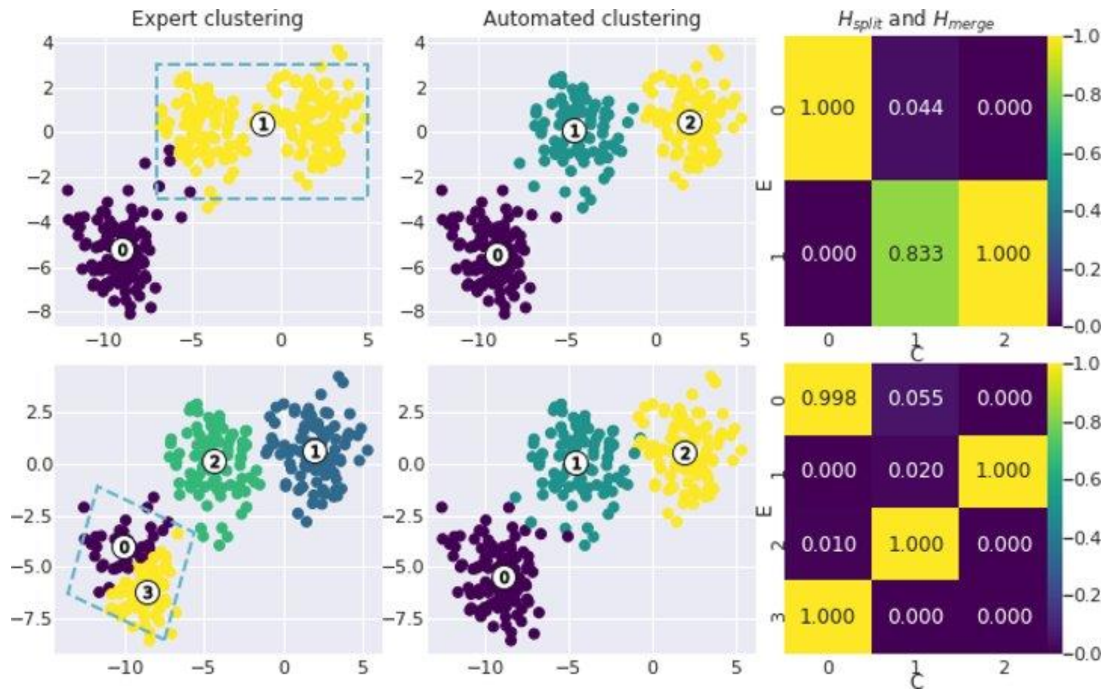
Expert clusters 2 and 4 are simmilar in their distribution in automated clustering

Merging expert cluster

- We calculated *l2* normalized distribution matrix
- We calculated cosine similarity between rows to denote expert clusters that were similarly splitted with automated method



Results - splits



$$C_i^{split} = \left\{ c_j \in H_i^{split} : \frac{c_j}{1-\lambda^s} > \varepsilon_s \right\}$$

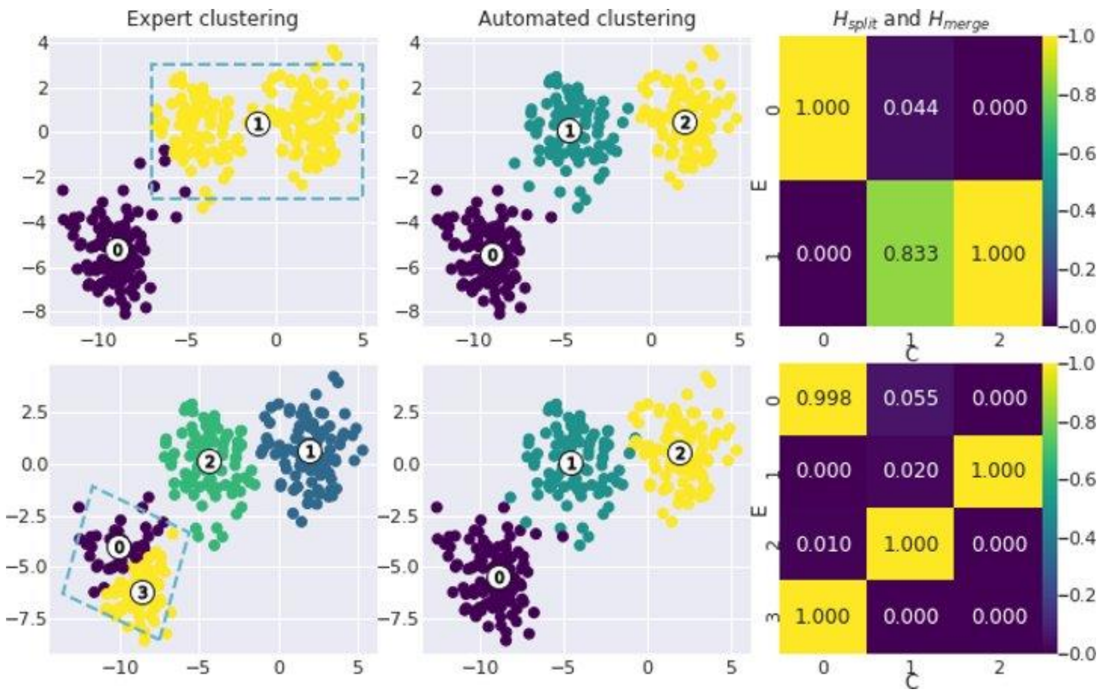
Decrease in silhouette score between splitted clusters

$$Conf(C_i^{split}) = \left\{ (1-\lambda^s)c_j + \lambda^s (S^{dec}(C_i^{split})) : c_j \in C_i^{split} \right\}$$

Assuming $\lambda^s = 0.1$

```
SPLIT EXPERT CLUSTER
  E_1
INTO CLUSTERS
  [(C_1, C_2)]
(Confidence 0.87)
```

Results - merges



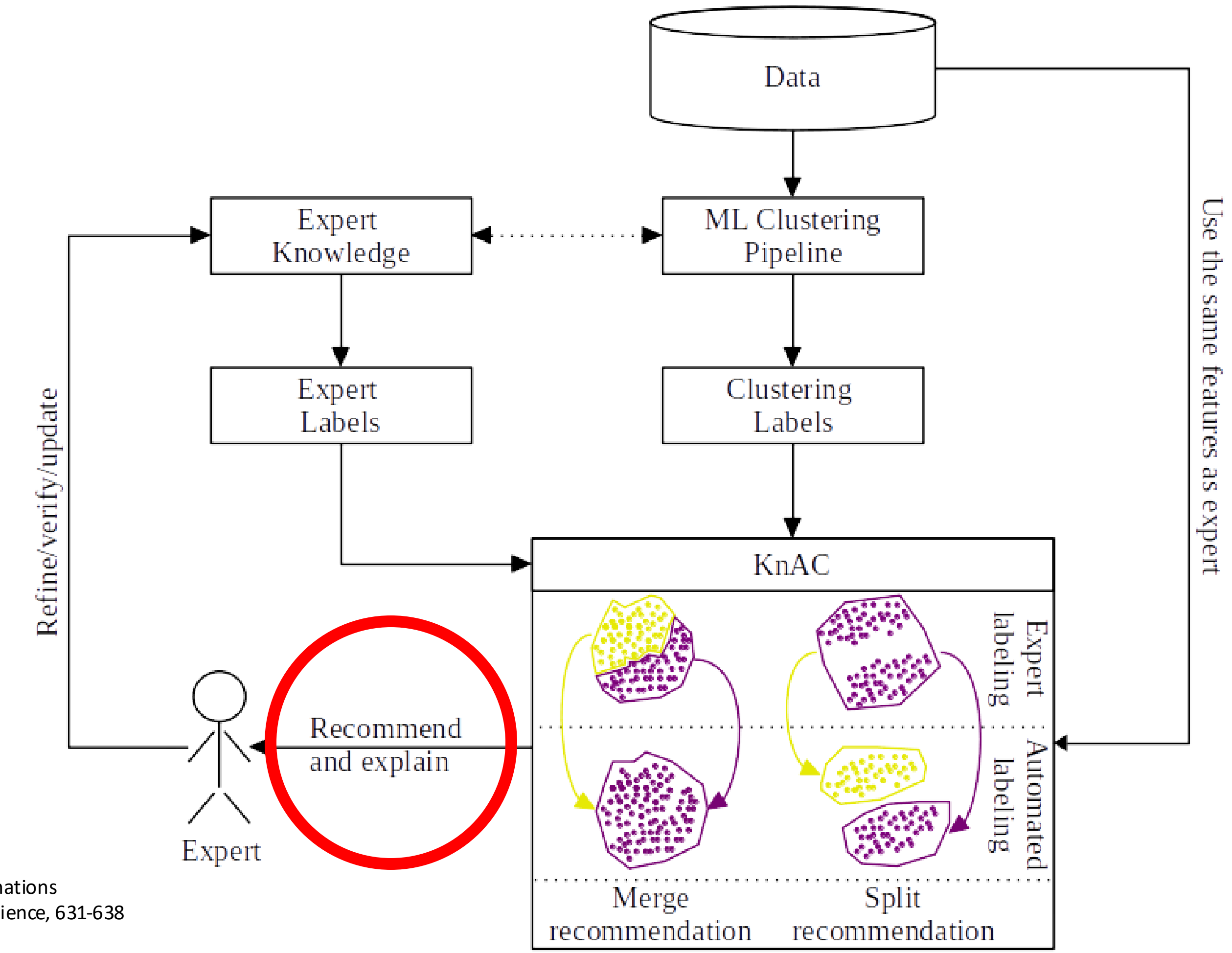
Linkage distance between merged clusters

$$C_{j,k}^{merge} = \left\{ E \ni E_j, E_k : (1-\lambda^m) H_{j,k}^{sim} + \lambda^m (1 - D_{j,k}^{linkage}) > \varepsilon_m \right\}$$

Assuming $\lambda^m = 0.2$

MERGE
EXPERT CLUSTER E_0
WITH
EXPERT CLUSTER E_3
INTO
CLUSTER C_0 # (Confidence 0.98)

Knowledge Augmented Clustering (KnAC)



Augmenting Automatic Clustering with Expert Knowledge and Explanations
S. Bobek, G.J. Nalepa, International Conference on Computational Science, 631-638

KnAC: an approach for enhancing cluster analysis with background knowledge and explanations, S. Bobek, M. Kuk, J. Brzegowski, E. Brzychczy, and G. J. Nalepa.
ArXiv: https://arxiv.org/abs/2112.08759

https://github.com/sbobek/knac

Why XAI is non trivial

*In an **act** of explaining, **someone** who is in possession of **some information***

Artificial intelligence

Feature contribution

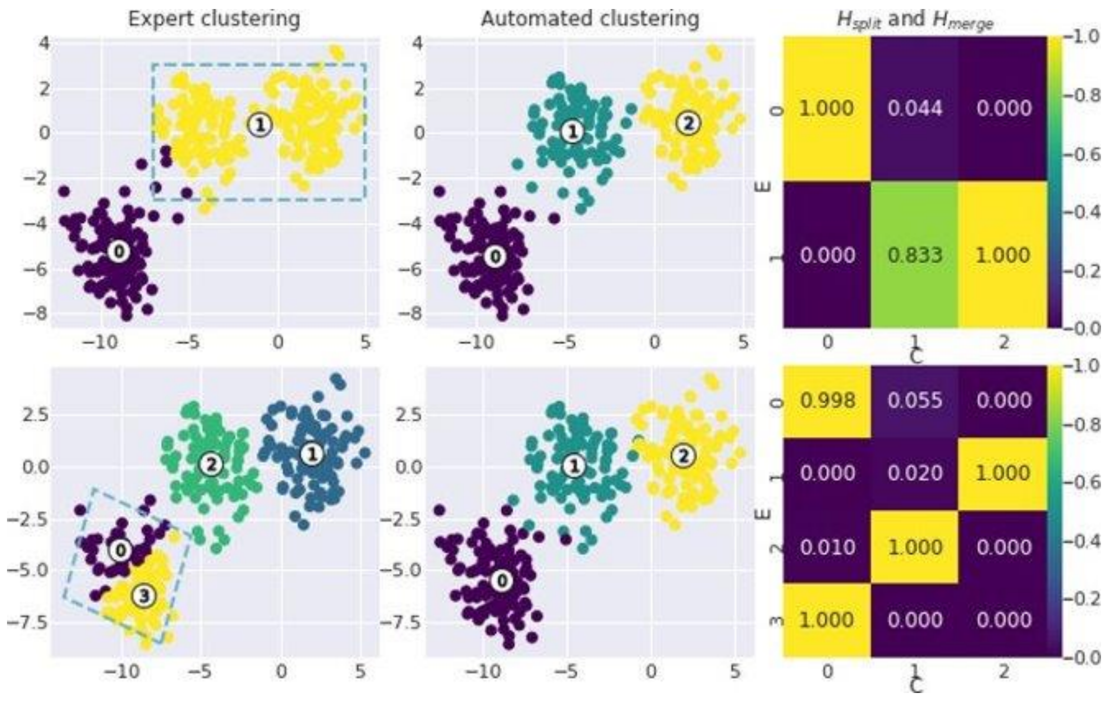
*about the **causal history of some event** - explanatory information,*

Why input to the model generated
such output

*I shall call it - tries to **convey it to someone else**.*

Human

How to use XAI in KnAC?



- Split: What makes the two new clusters different from each other to convince expert they are different entities?
- Merge: What makes the two expert clusters different from each other to convince expert that they are the same entity (difference is irrelevant)

From clustering to classification

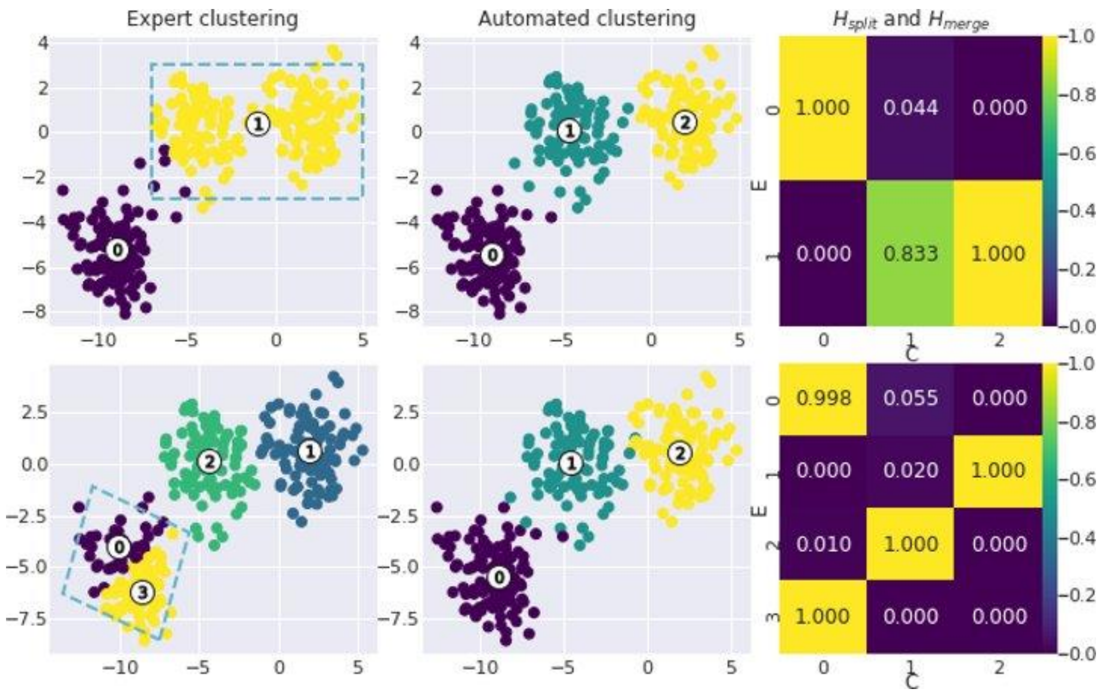
Data
- Transformations
- Feature extraction

Clustering
- Deciding on number of clusters
- No limits on the clustering algorithm

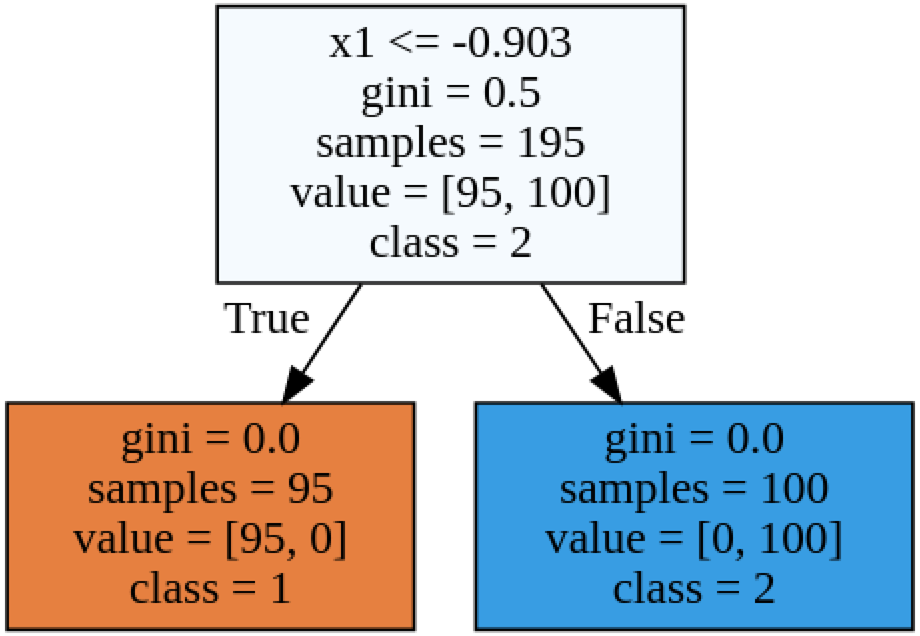
Classification
- Transform clustering into classification
- Use cluster labels as target class
- Use features understandable by the user

Explanation
- Apply XAI methods to obtain reason why two clusters are different

Explanations of splits

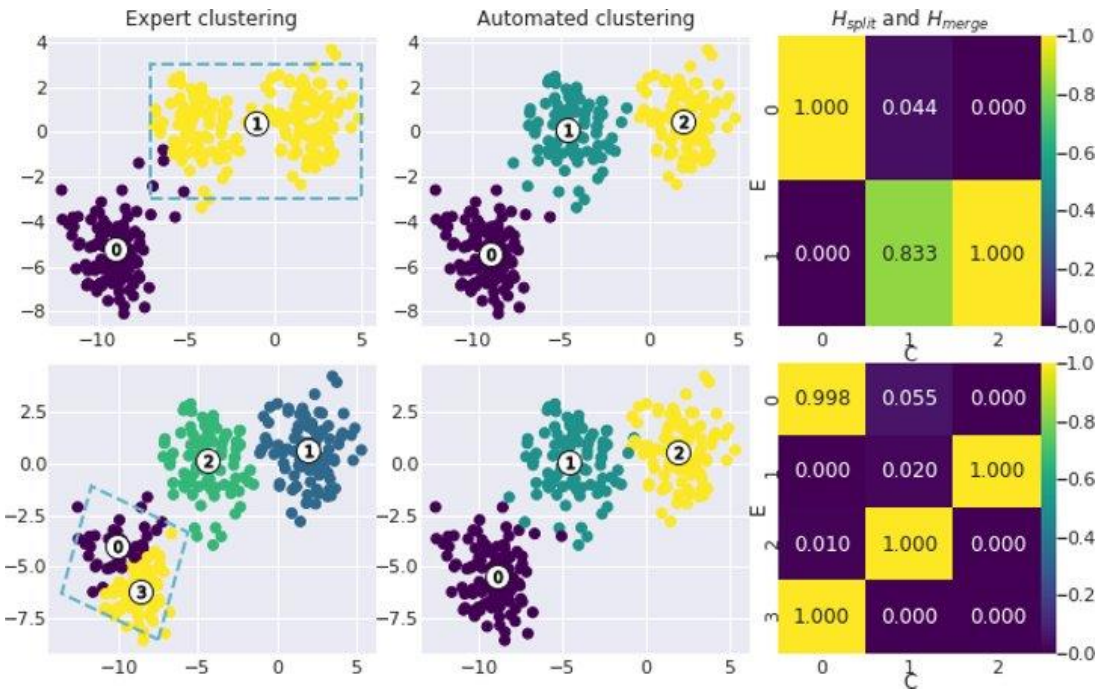



```
SPLIT EXPERT CLUSTER
  E_1
INTO CLUSTERS
  [(C_1, C_2)]
  (Confidence 0.87)
```

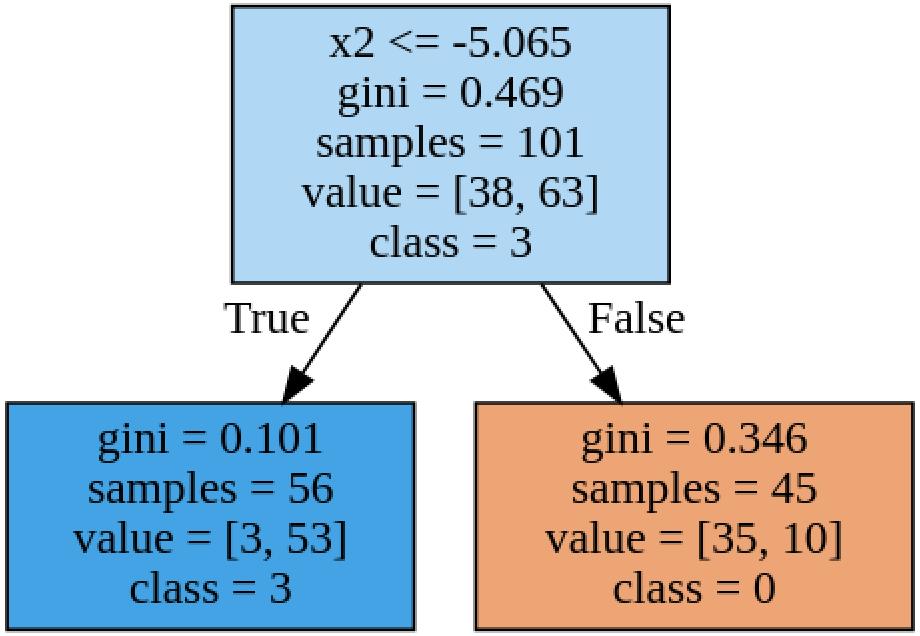



```
C_1: x1 <= -0.30 (Precision: 0.99, Coverage: 0.49)
C_2: x1 > -0.30 (Precision: 1.00, Coverage: 0.49)
```

Explanations of merges

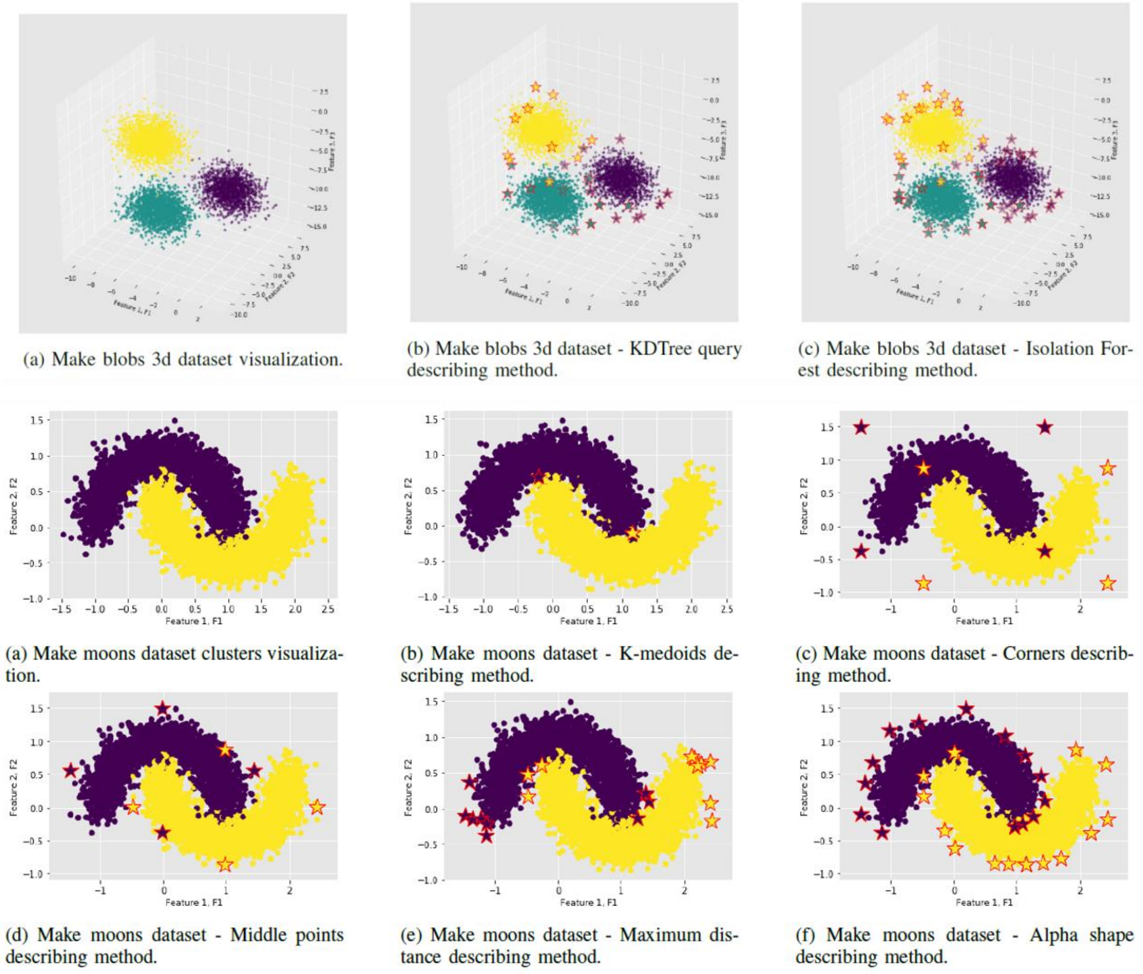



```
MERGE
    EXPERT CLUSTER E_0
WITH
    EXPERT CLUSTER E_3
INTO
    CLUSTER C_0
    (Confidence 0.98)
```




```
E_0: x1 <= -8.20 AND x2 > -4.34 (Precision: 1.00, Coverage: 0.07)
E_1: x1 <=-4.34 (Precision: 0.90, Coverage: 0.25)
```

Explainable clusters



(a) Make blobs 3d dataset visualization.

(b) Make blobs 3d dataset - KDTree query describing method.

(c) Make blobs 3d dataset - Isolation Forest describing method.

(a) Make moons dataset clusters visualization.

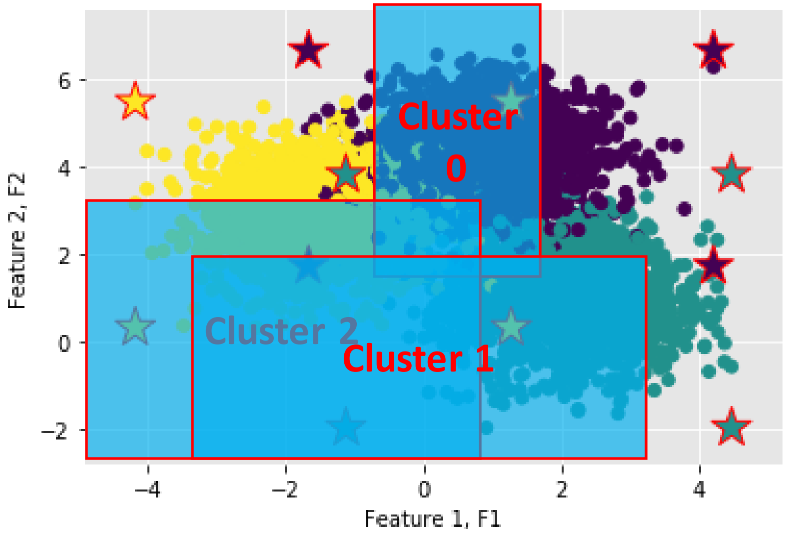
(b) Make moons dataset - K-medoids describing method.

(c) Make moons dataset - Corners describing method.

(d) Make moons dataset - Middle points describing method.

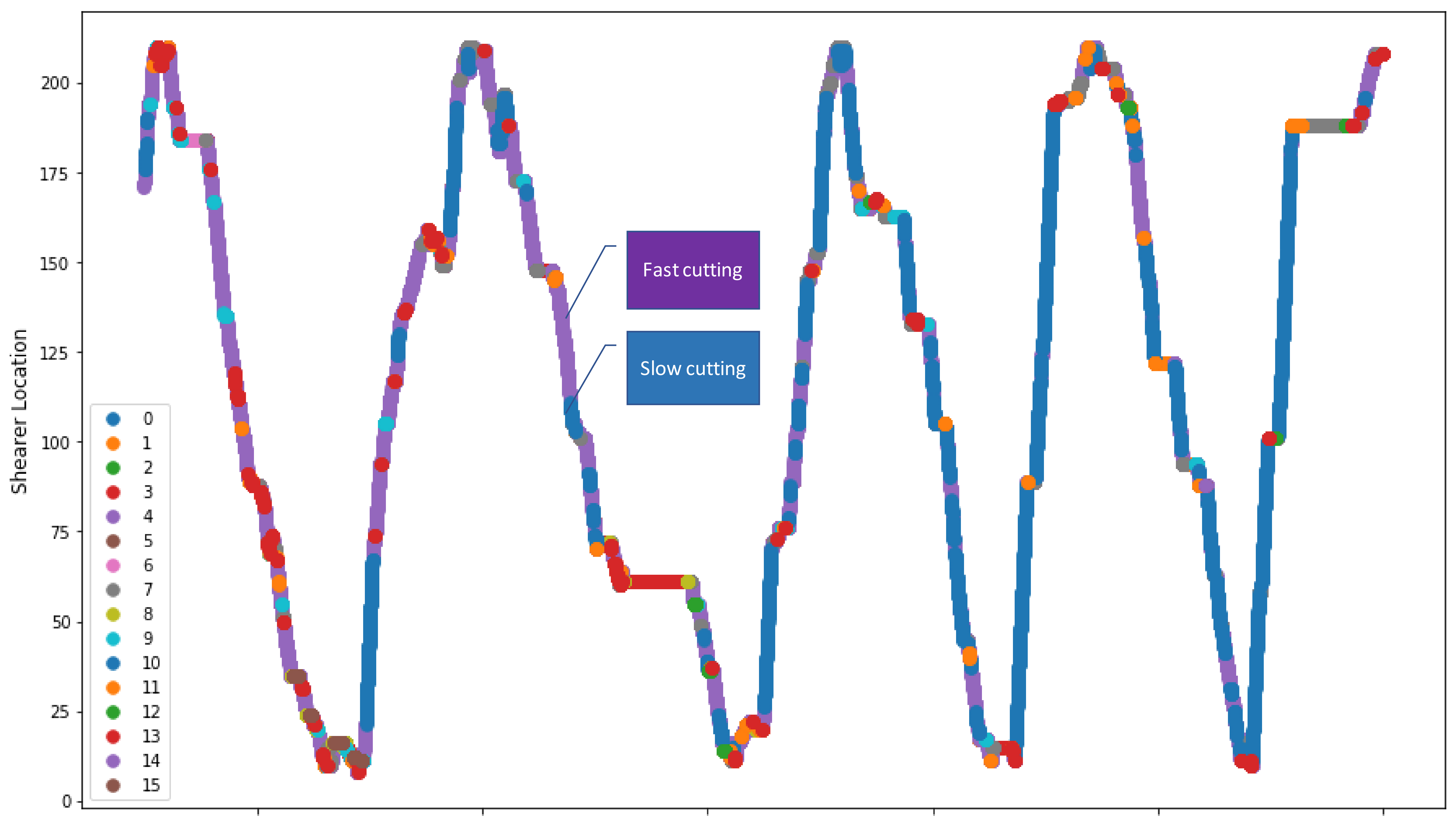
(e) Make moons dataset - Maximum distance describing method.

(f) Make moons dataset - Alpha shape describing method.



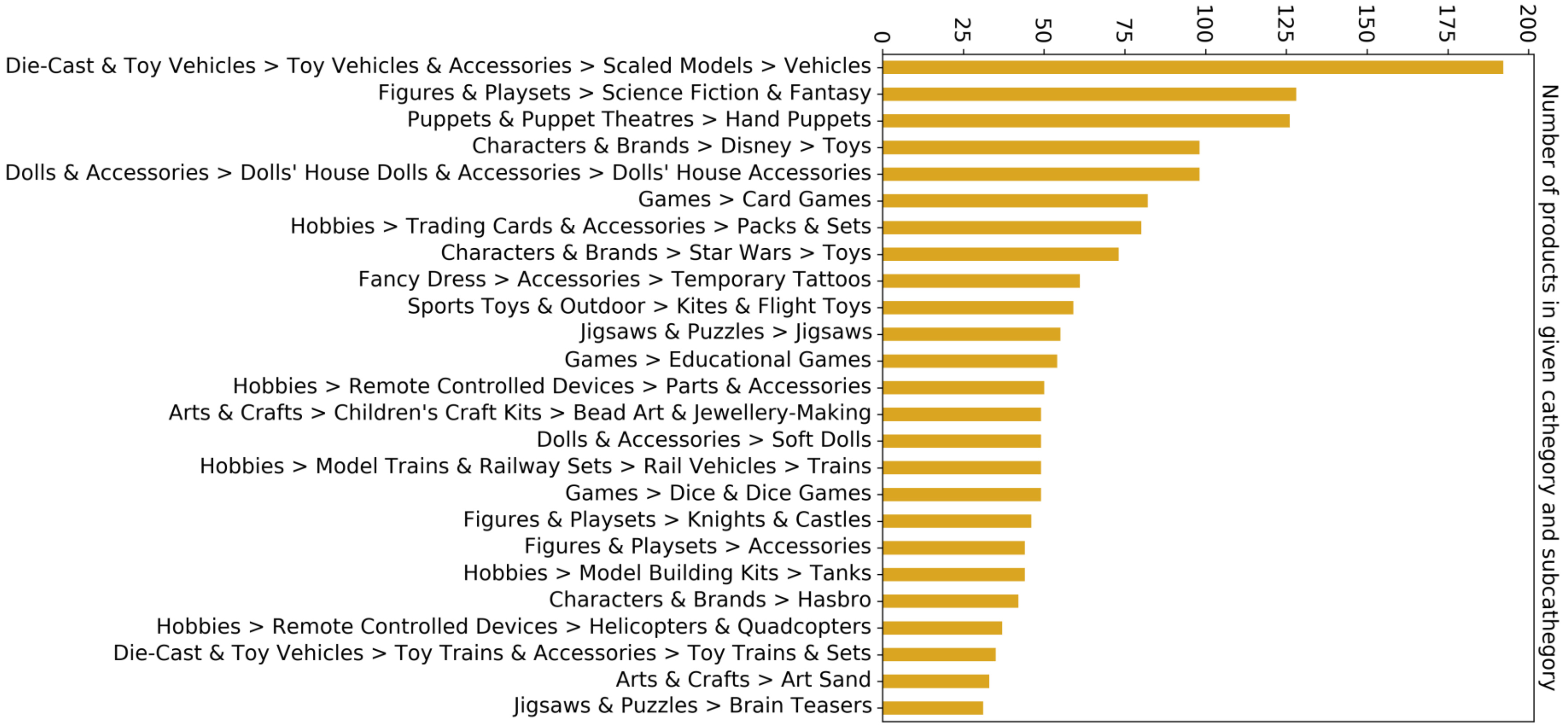
Rule no.	Rule	Cluster	Certainty
1	F1 > 0.68 and F2 > 2.99	0	0.48
2	0.68 < F1 ≤ 1.77 and F2 > 1.64	0	0.64
3	-1.14 < F1 ≤ 1.77 and F2 > 1.64	0	0.54
4	F1 > 0.68 and F2 ≤ 2.99	1	0.44
5	F1 > -1.14 and F2 ≤ 1.64	1	0.68
6	F1 ≤ -1.14	2	0.25
7	F1 ≤ 0.68 and F2 ≤ 2.99	2	0.43

M. Kuk, S. Bobek and G. J. Nalepa, "Explainable clustering with multidimensional bounding boxes," *2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA)*, 2021, pp. 1-10, doi: 10.1109/DSAA53316.2021.9564220.



Shearer Location
200
175
150
125
100
75
50
25
0
Fast cutting
Slow cutting
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

E-commerce and coal mine



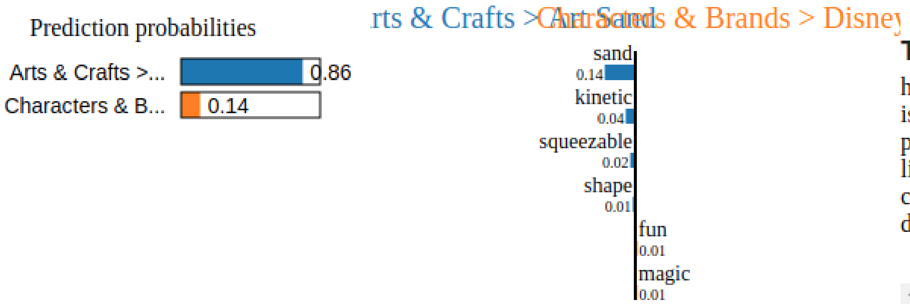
Product to category

Chuggington is an action-packed contemporary animated train series for pre-schoolers that follows the exciting adventures of three young trainees: Wilson, Brewster and Koko. In each energetic, vibrant episode, the trainees ride the rails through the world of Chuggington, exploring many locations and taking on exciting challenges that test their courage, speed and determination. With the help support and guidance of the more experienced Chuggers, they learn positive values, including respect and loyalty, and new skills such as teamwork and patience, empowering them to be the best trainees they can be. Box Contains 1 x Chuggington Train ”

E-commerce and coal mine

See the results at online tutorial: **https://github.com/sbobek/knac**

```
Justification for Arts & Crafts > Art Sand
[('sand', -0.14062779721924024), ('kinetic', -0.04427780308500065), ('squeezable', -0.023407082689875347), ('shape', -0.011500639821694475), ('fun', 0.0103043090148735
```



Text with highlighted words

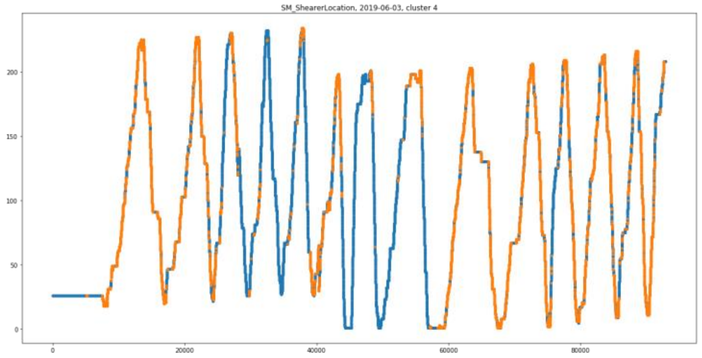
have hours of fun with this magic sand playset creating brilliant sand shapes or create your own sculptures the magic sand is the squeezable sand where you can feel the fun pack it pull it shape it and love it motion sand is so incredible you can put it down it kinetic meaning it sticks to itself and not to you it easy to shape and mould and flows through your fingers like slow moving liquid but leaves them completely dry kinetic sand stimulates children creative skills allowing them to create anything they can imagine it never dries out and is gluten free this soft and stretchy sand easily cleans up while delivering non stop fun it squeezable sand you can put down for ages years and over

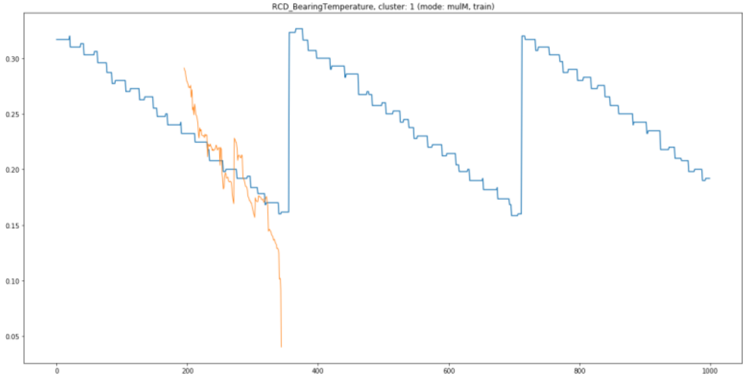
```
Justification for Characters & Brands > Disney > Toys
[('disney', 0.027092040859663918), ('soft', -0.008068585098757642), ('character', 0.006040355286770795), ('characters', 0.004645787316197593), ('kids', 0.0043490668605
```

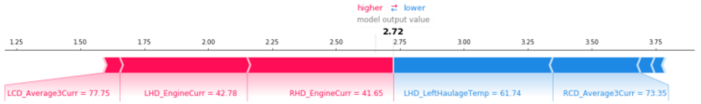
Text with highlighted words

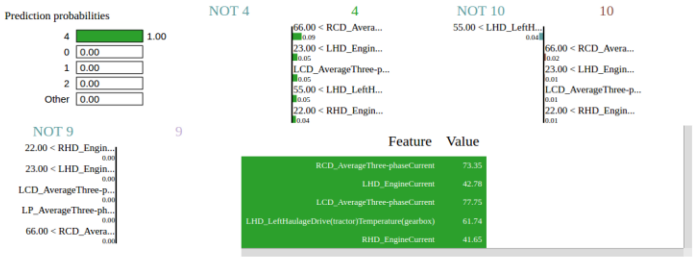
product description whether at home on the road or in the air your favourite disney character can provide great companyand comfort these soft colourful cushions can be easily transformed into disney character soft toy by simply opening and closing the velcro loved by children of all ages these classic disney characters will keep kids entertained for hours and when sleepy just rest your head on the cushion and dream away all our disney character cushions are washable please read washing label for further instructions box contains x

How to explain? Which explanation we should trust?



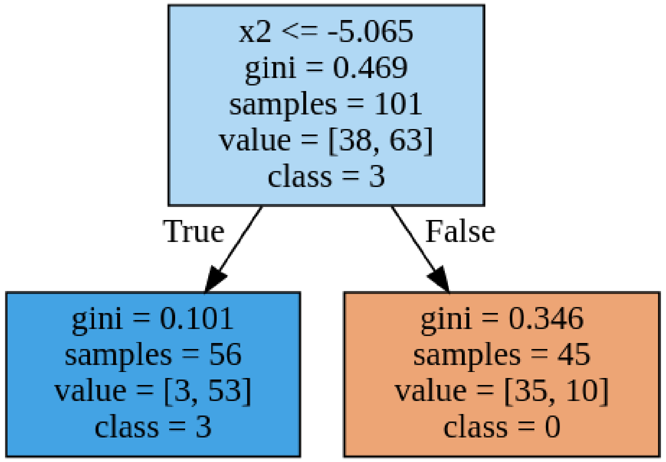




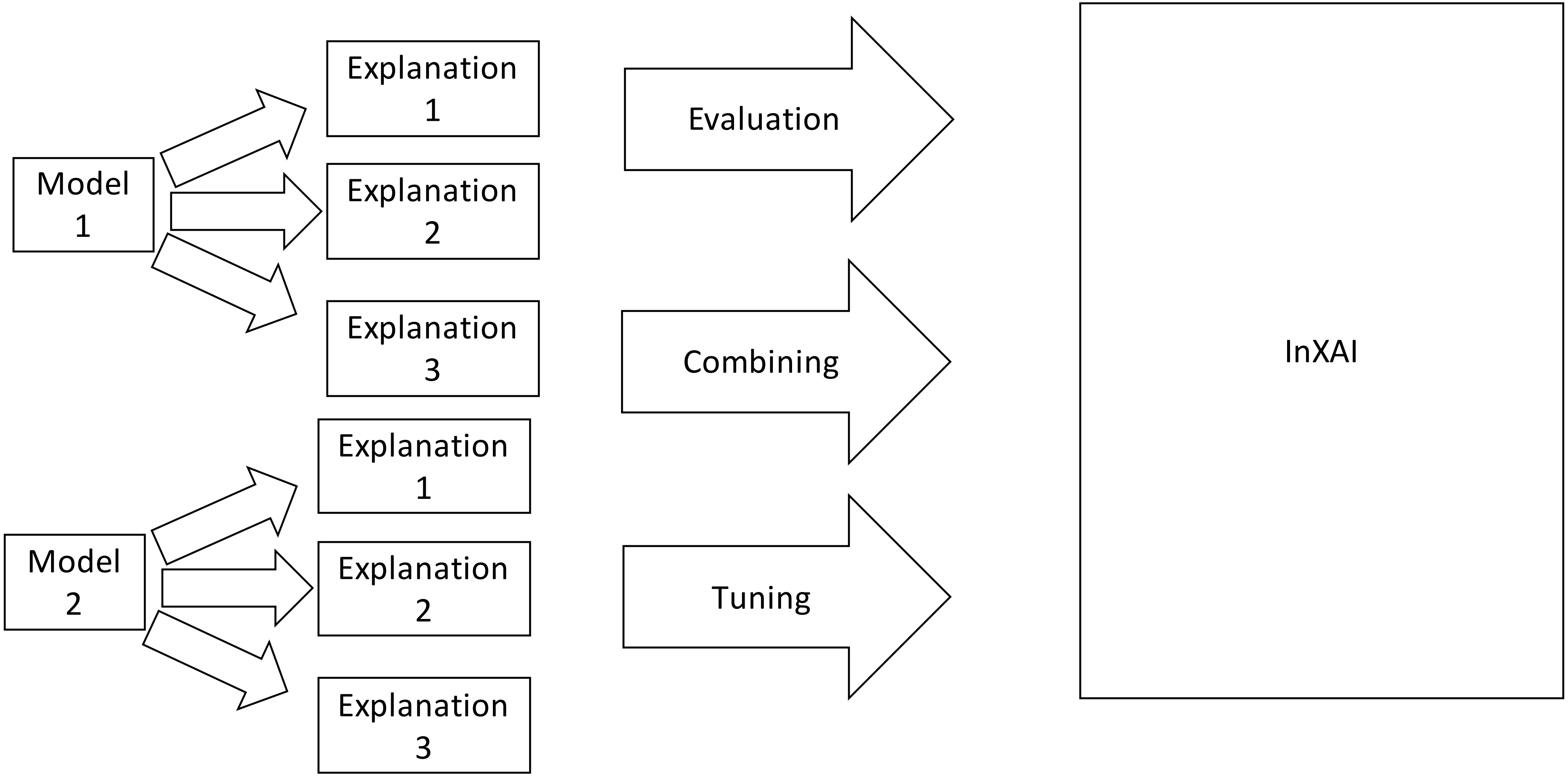


ANCHOR

SM_ShearerSpeed <= 2.00 AND
(RCD_AverageThree-phaseCurrent_LOW OR
RCD_AverageThree-phaseCurrent_MEDIUM OR
RCD_AverageThree-phaseCurrent_HIGH OR
RCD_AverageThree-phaseCurrent_OVERLOAD) AND
(RHD_EngineCurrent_IDLE OR
RHD_EngineCurrent_LOW) AND
(LCD_AverageThree-phaseCurrent_IDLE OR
LCD_AverageThree-phaseCurrent_LOW) AND
RHD_RightHaulageDrive(tractor)Temperature(gearbox) > 0.00 AND
(LHD_EngineCurrent_IDLE OR LHD_EngineCurrent_LOW) AND
43.00 < LA_LeftArmTemperature <= 52.00 AND
LHD_LeftHaulageDrive(tractor)Temperature(gearbox) > 65.00 AND
RA_RightArmTemperature <= 54.00 AND LP_AverageThree-phaseCurrent
<= 4.00 AND
57.00 < RCD_BearingTemperature <= 64.00



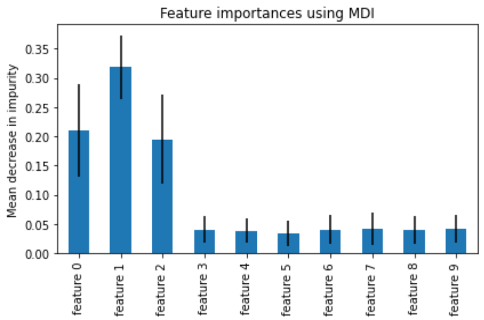
Intelligible XAI (InXAI)

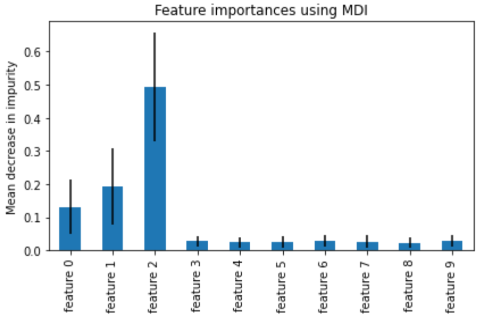


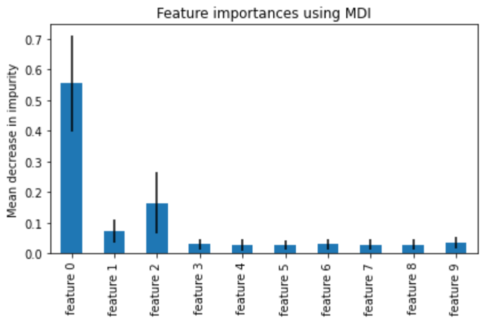
https://github.com/sbobek/inxai

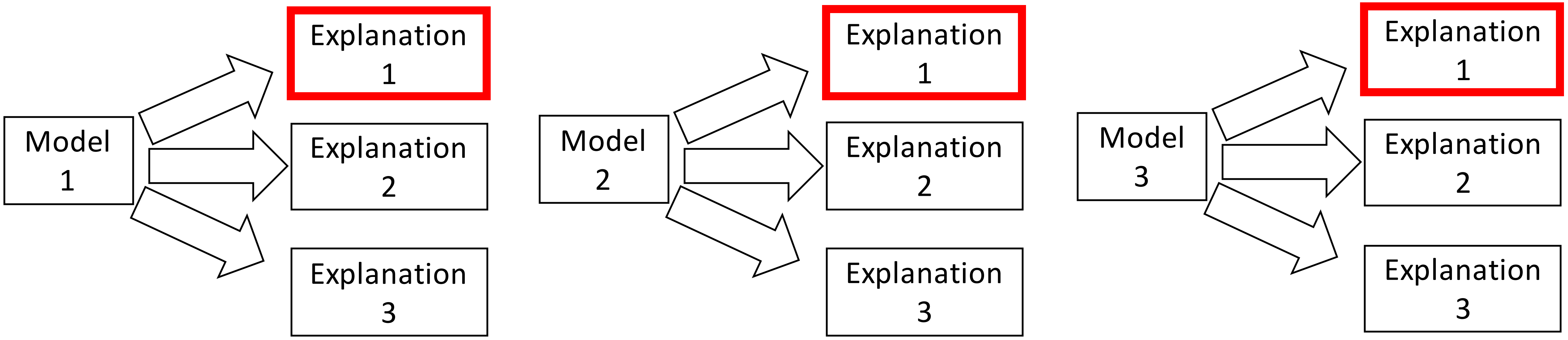
Consistency between explanations for diferent models (or explainers)

$$C(\Phi^{e \to m_1}, \Phi^{e \to m_2}, \ldots, \Phi^{e \to m_n}) = \frac{1}{\max\limits_{a,b \in m_1, m_2, \ldots, m_n} ||\Phi_j^{e \to m_a} - \Phi_j^{e \to m_b}||_2 + 1}$$



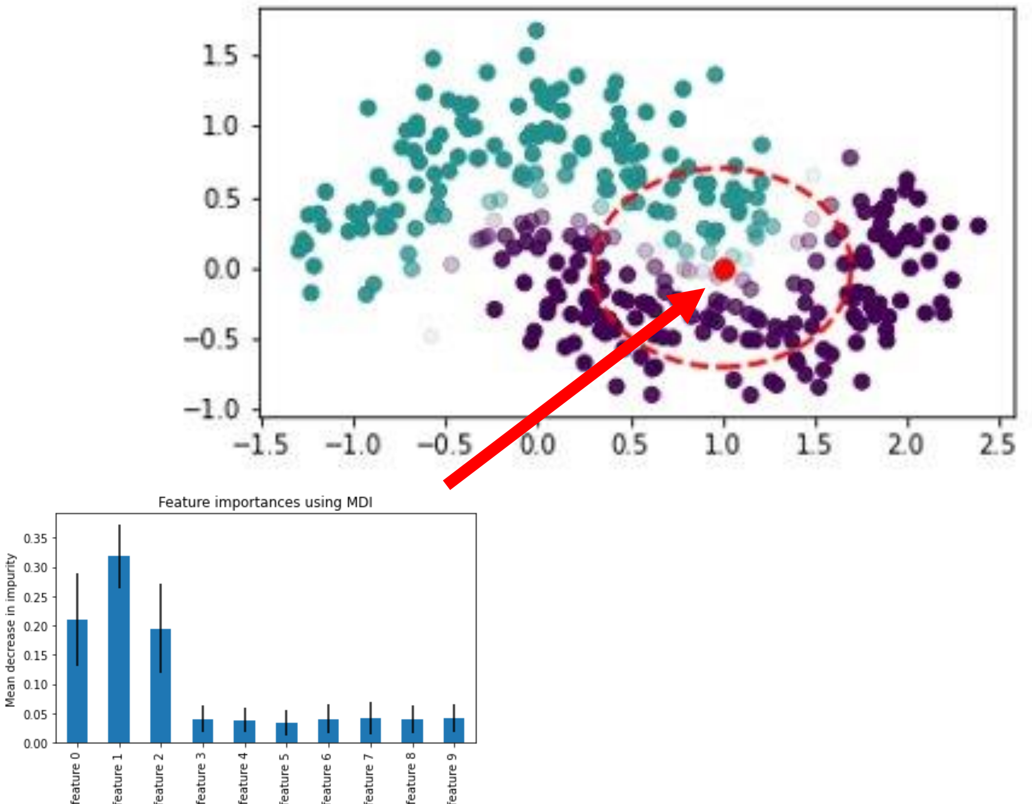


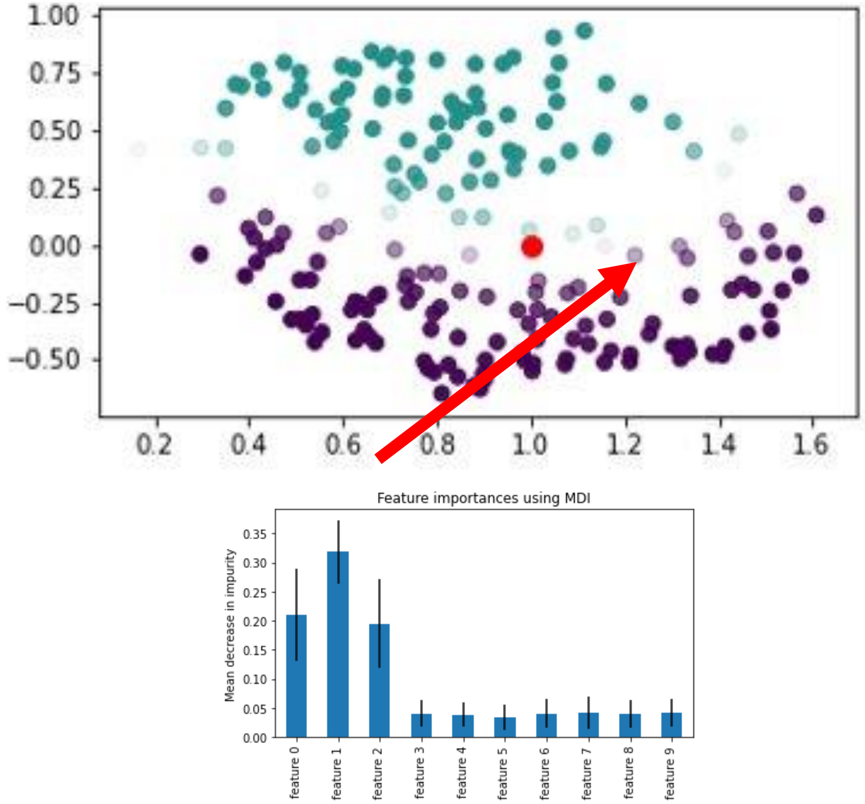




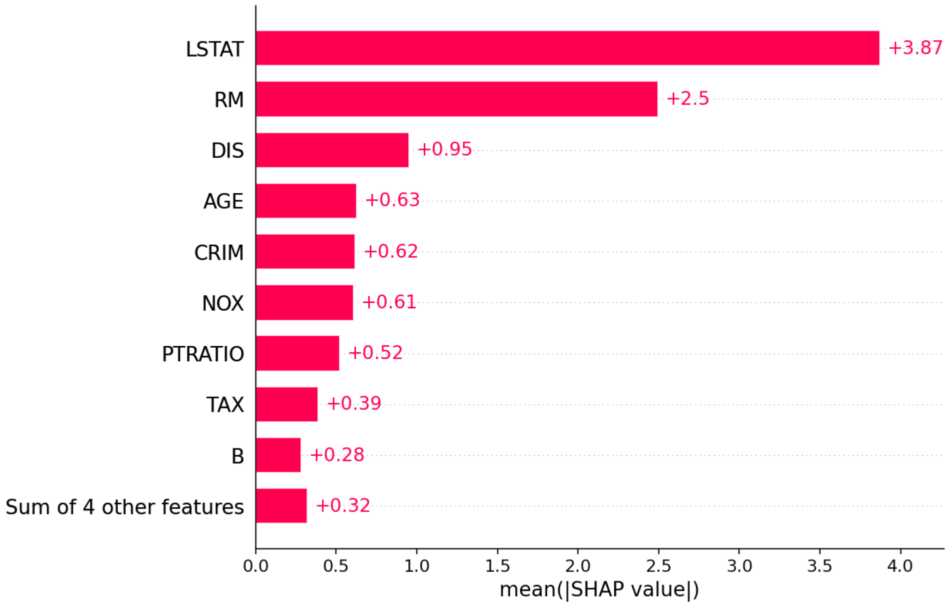
Stability of explanations for similar instances

$$\hat{L}(\Phi^{e\rightarrow m}, X) = \max_{x_j \in N_\epsilon(x_i)} \frac{||x_i - x_j||_2}{||\Phi_i^{e\rightarrow m} - \Phi_j^{e\rightarrow m}||_2 + 1}$$

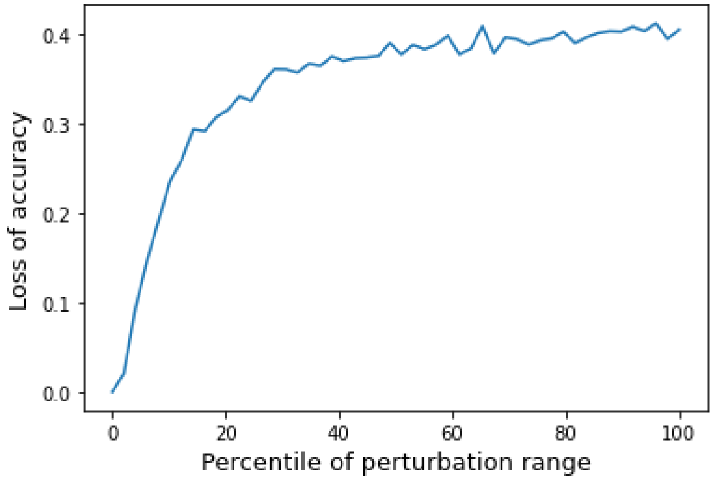


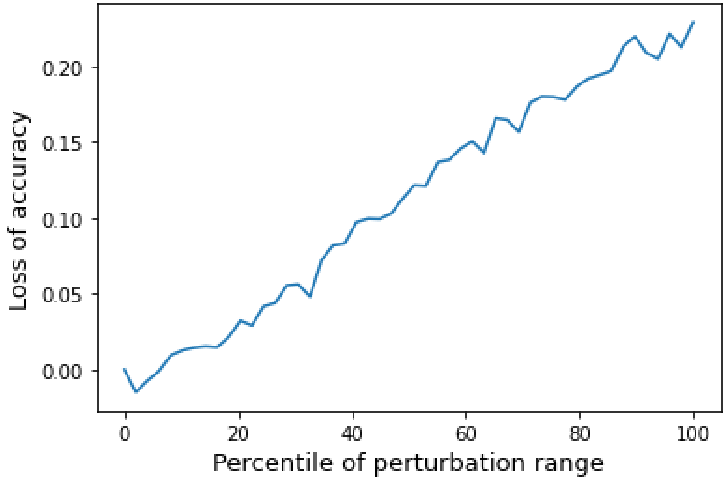


Quality Loss (AUCx)

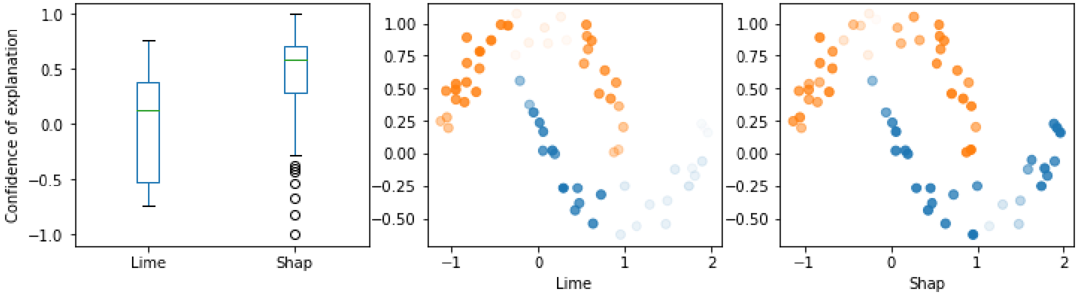


Perturb data accordingly





Ensemble explanations

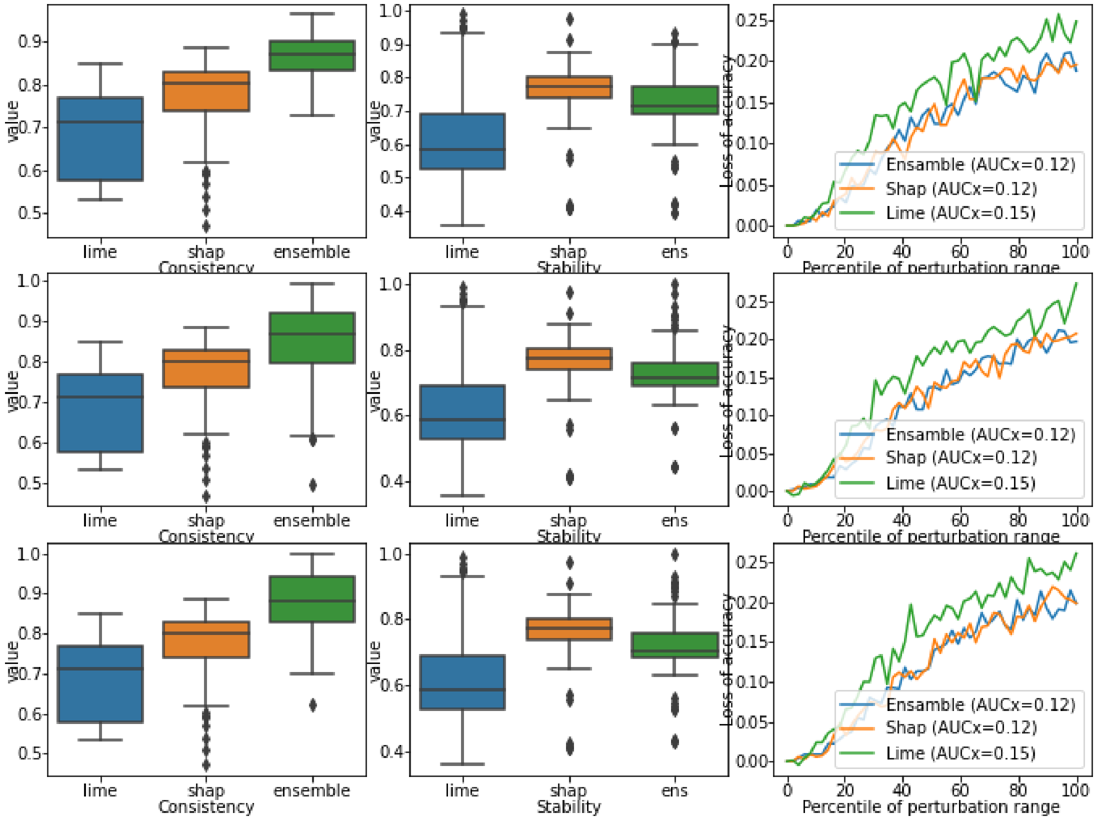


$$C(\Phi^{e\to m_1}, \Phi^{e\to m_2}, \dots, \Phi^{e\to m_n}) = \frac{1}{\max\limits_{a,b\in m_1,m_2,\dots,m_n} ||\Phi_j^{e\to m_a} - \Phi_j^{e\to m_b}||_2 + 1}$$

$$\hat{L}(\Phi^{e\to m}, X) = \max_{x_j \in N_\epsilon(x_i)} \frac{||x_i - x_j||_2}{||\Phi_i^{e\to m} - \Phi_j^{e\to m}||_2 + 1}$$

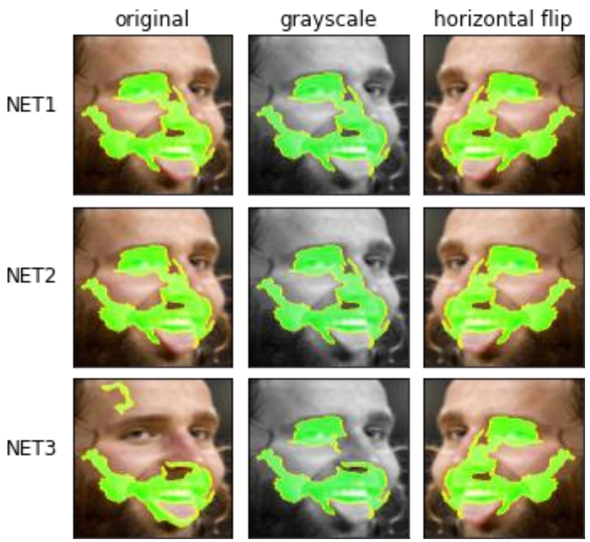
$$ES(M, w) = \sum w_i \cdot M_i$$

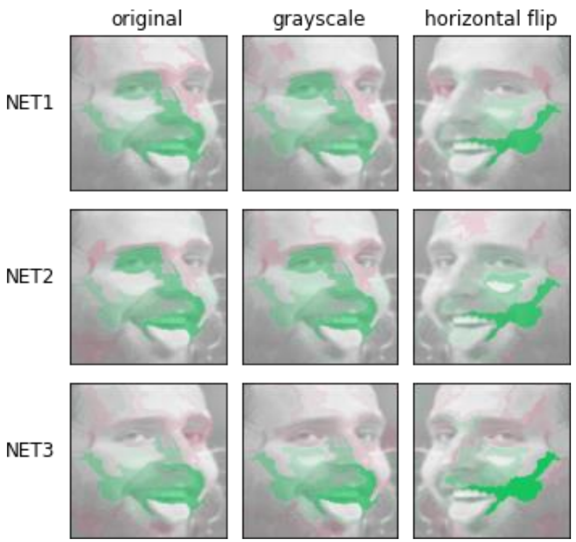
$$\Phi^{ens} = \frac{ES(M,w)\cdot[\gamma_1\Phi^{e_1}, \gamma_2\Phi^{e_2}, \dots, \gamma_n\Phi^{e_n}]}{\sum_{i=1}^{n} ES_i(M,w)}$$

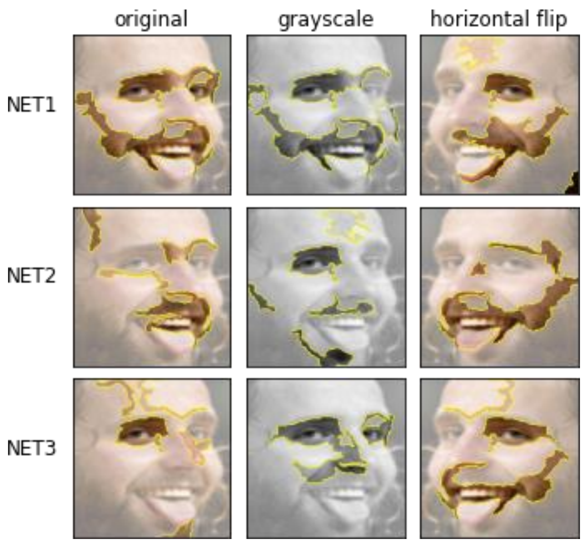


https://github.com/sbobek/inxai

Ensemble explanations

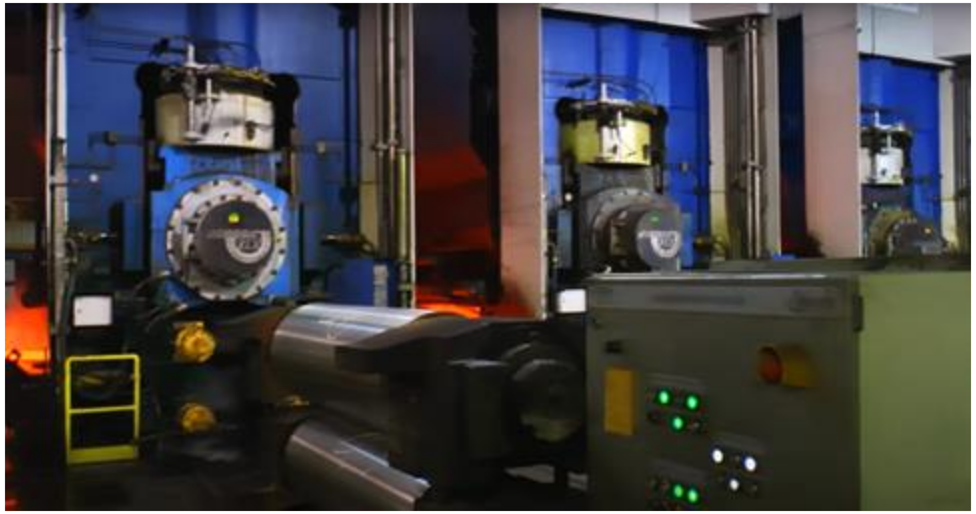






Anomaly detection

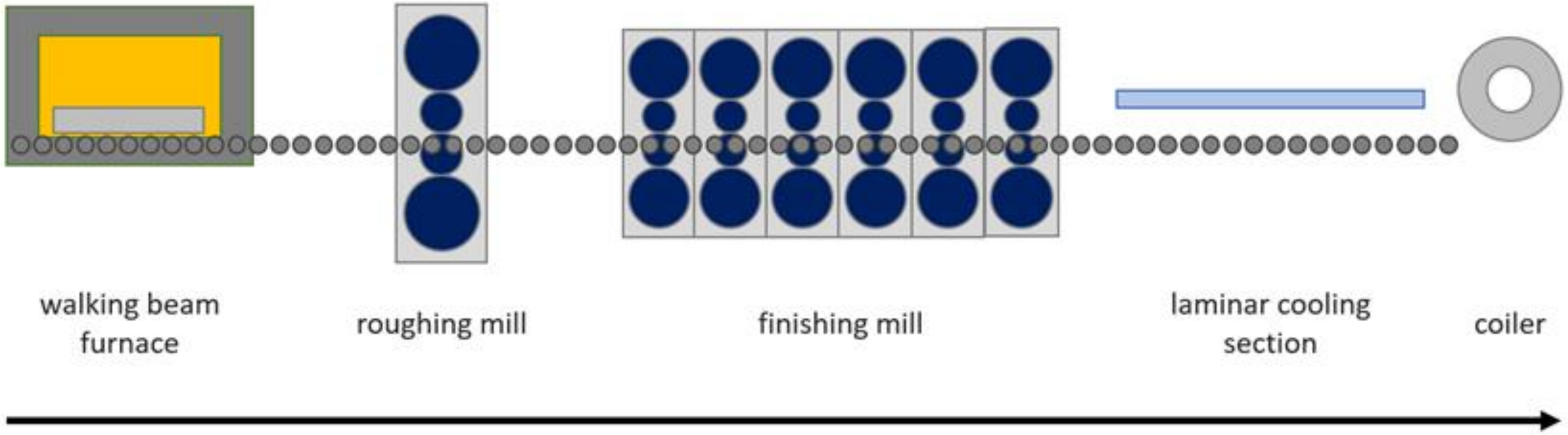
Anomaly detection in hot- and cold-rolling process



Anomaly detection in hot- and cold-rolling process

Hot-rolling process

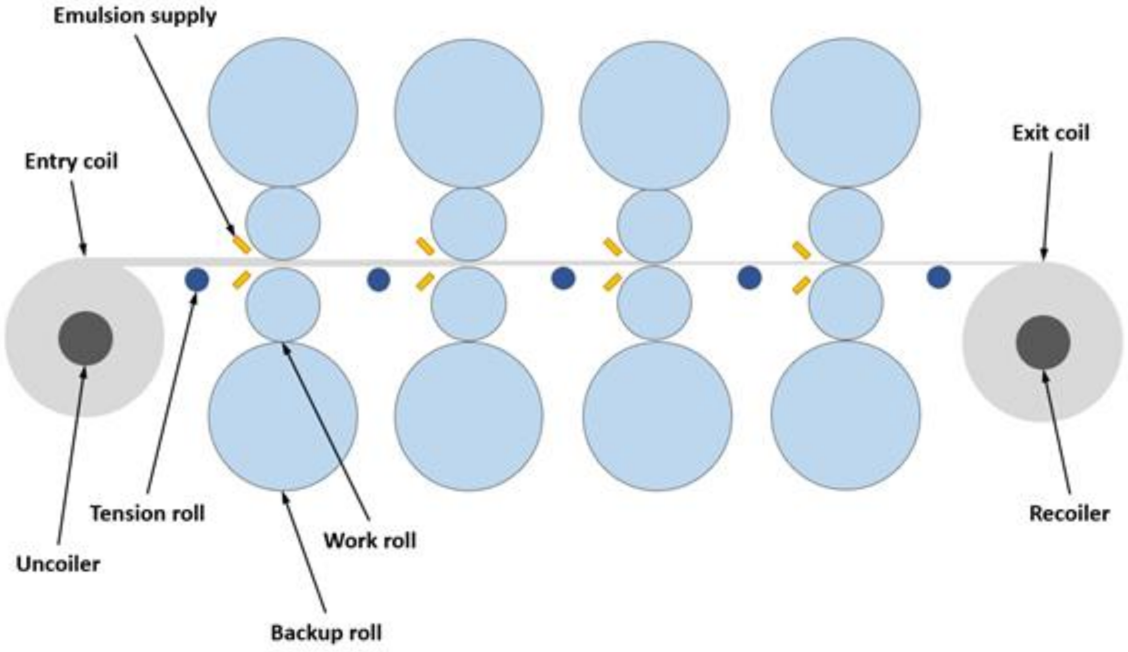
- After casting a steel slab has dimensions of approximately 0.22m x 1.5m x 10m.
- Further processing is needed to obtain the shape and dimensions required by the clients.



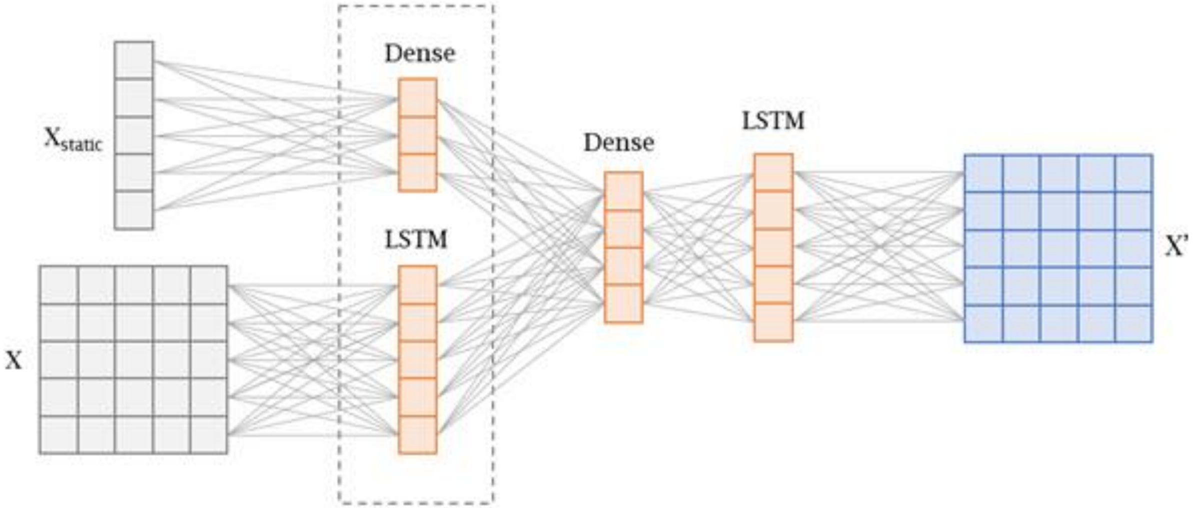
Anomaly detection in hot- and cold-rolling process

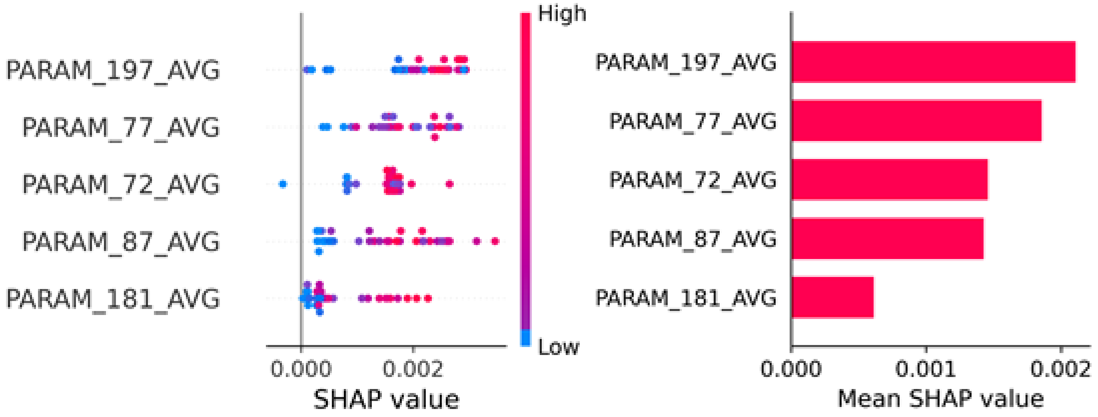
Cold-rolling process

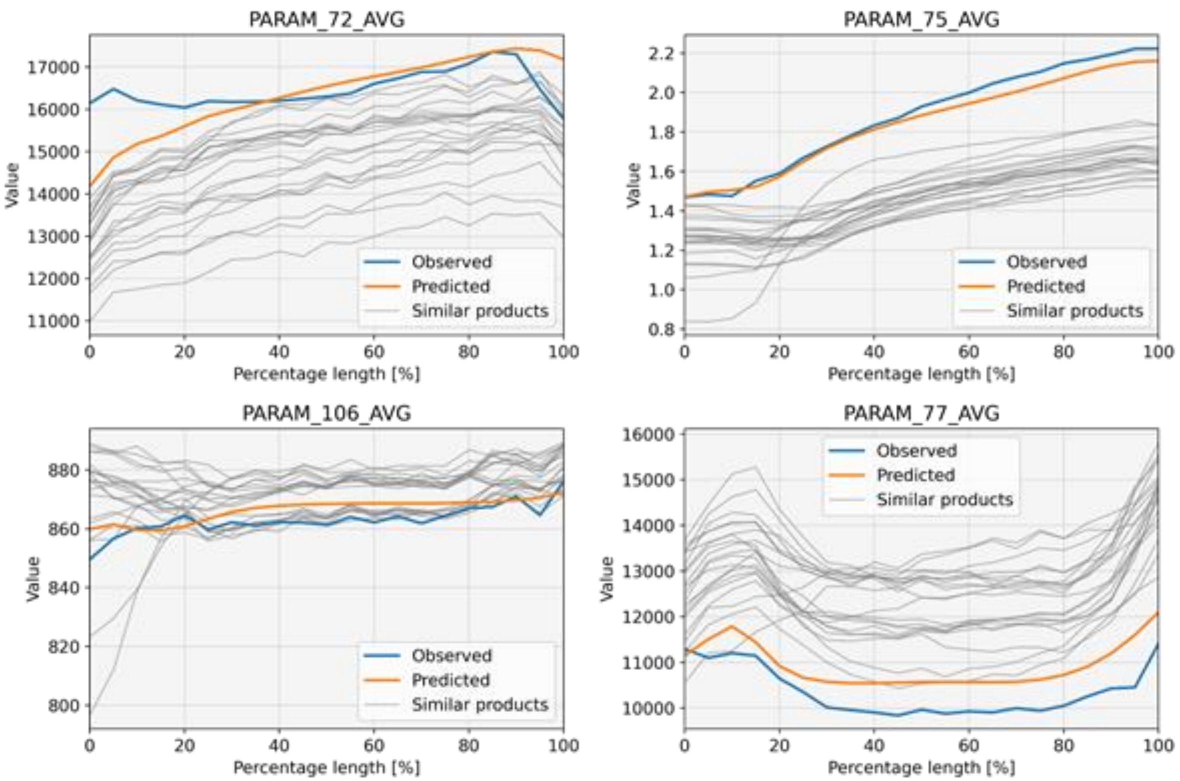
- Optional production step after hot rolling.
- Used to reduce the thickness of steel strip without preheating by 30 to 80%.
- Use-case production line consists of four stands, which reduce



Anomaly detection in hot-rolling process



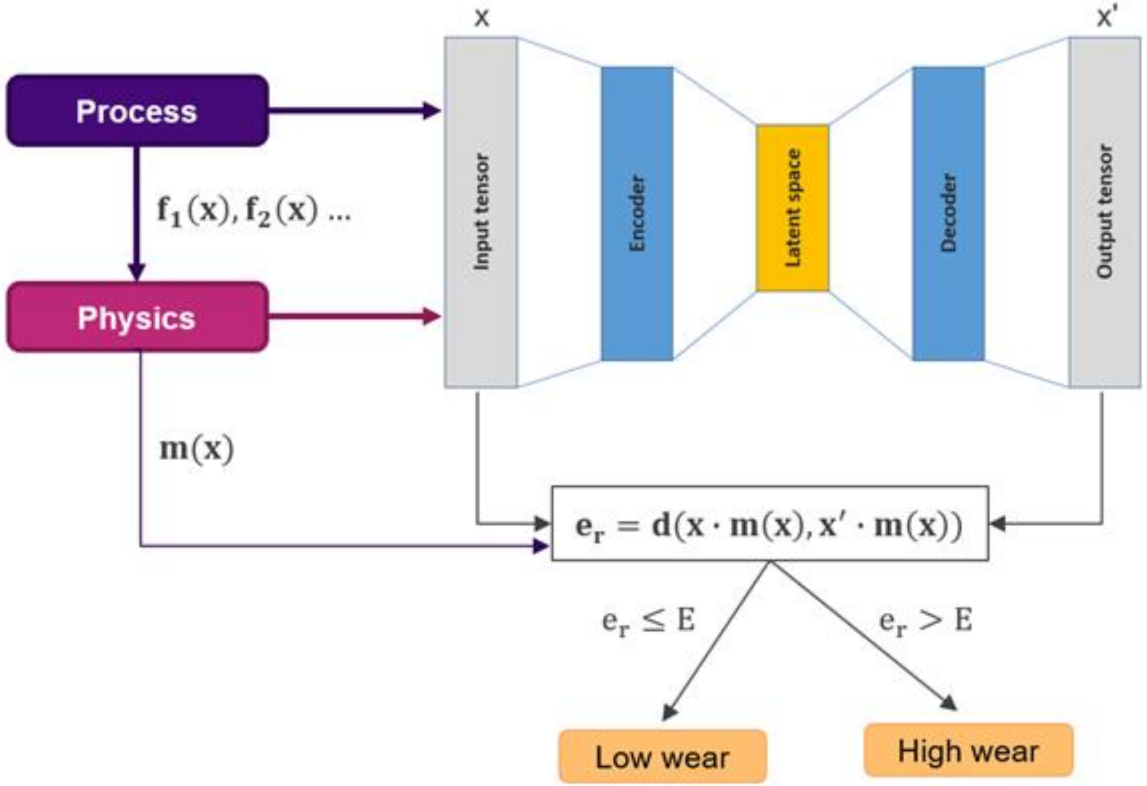




Jakubowski J, Stanisz P, Bobek S, Nalepa GJ. **Anomaly Detection in Asset Degradation Process Using Variational Autoencoder and Explanations**. *Sensors*. 2022; 22(1):291. https://doi.org/10.3390/s22010291

Anomaly detection in cold-rolling process

$$P_r = 2R' \int_0^{\phi_n} \frac{kh}{h_o}(1 - \frac{\sigma_o}{k_o}) \exp(\mu H)\, d\phi + 2R' \int_{\phi_n}^{\phi_o} \frac{kh}{h_i}(1 - \frac{\sigma_i}{k_i}) \exp(\mu(H_i - H))\, d\phi$$



Counterfactuals!

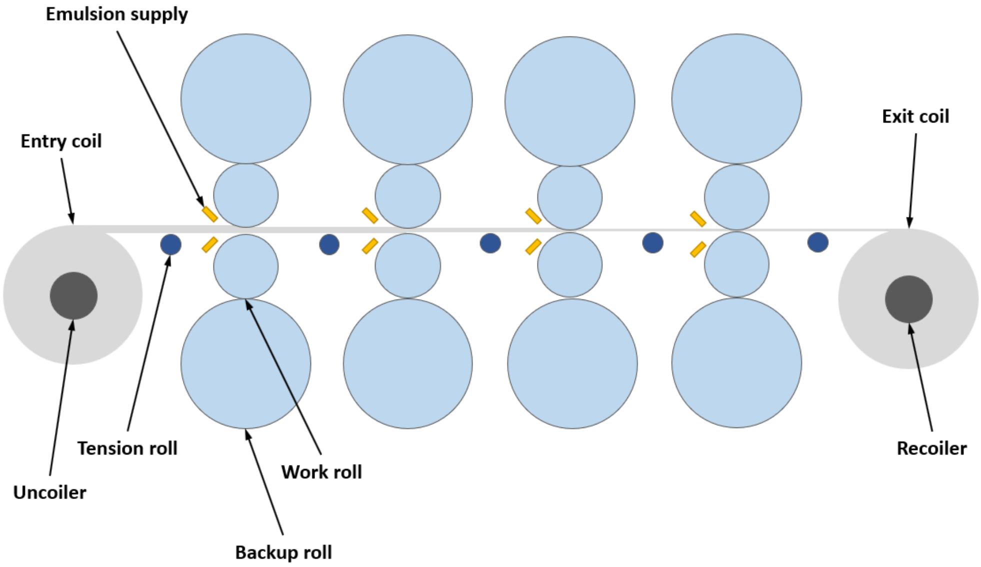
Metric	AE		PIAE	
	Test	Validation	Test	Validation
Accuracy	36.0%	31.2%	82.1%	79.9%
Precision	36.1%	30.2%	72.6%	61.1%
Recall	97.9%	100.0%	82.0%	88.9%
F1	52.7%	46.4%	77.0%	72.4%

J. Jakubowski, P. Stanisz, S. Bobek and G. J. Nalepa, "**Roll Wear Prediction in Strip Cold Rolling with Physics-Informed Autoencoder and Counterfactual Explanations**," *2022 IEEE 9th International Conference on Data Science and Advanced Analytics (DSAA)*, Shenzhen, China, 2022, pp. 1-10, doi: 10.1109/DSAA54385.2022.10032357.

Towards Online Anomaly Detection in Steel Manufacturing Process

The task is to detect anomalies in streaming data from cold rolling process. Several issues are addressed in this paper:

- product mix is heavily imbalanced (resampling)
- concept drift detection
- online learning algorithms are compared with batch learning equivalents



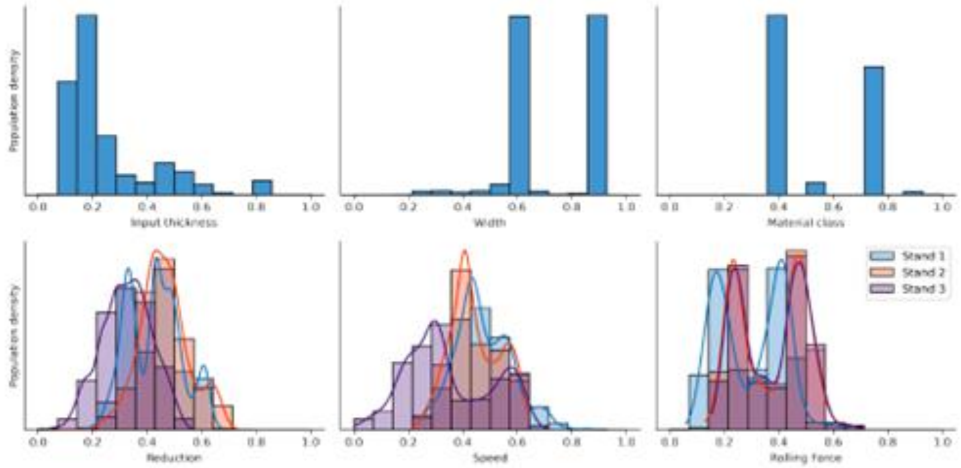


Fig. 2. Normalized data distribution of the most relevant features

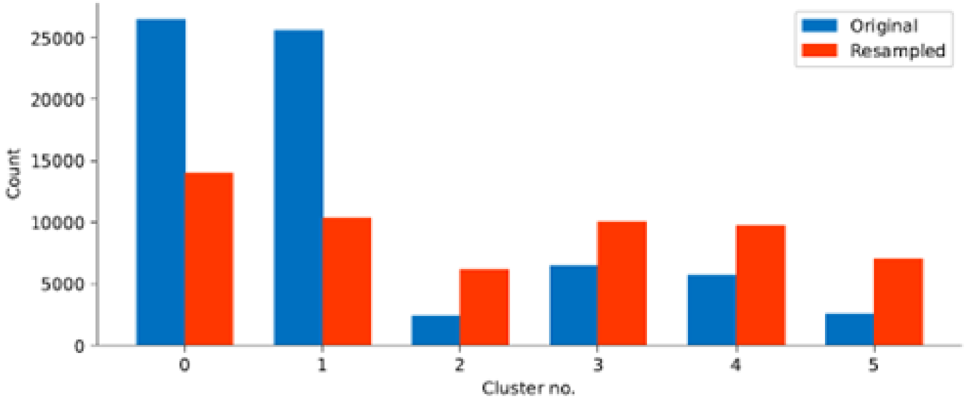


Fig. 3. Count of original and resampled data with respect to the assigned cluster.

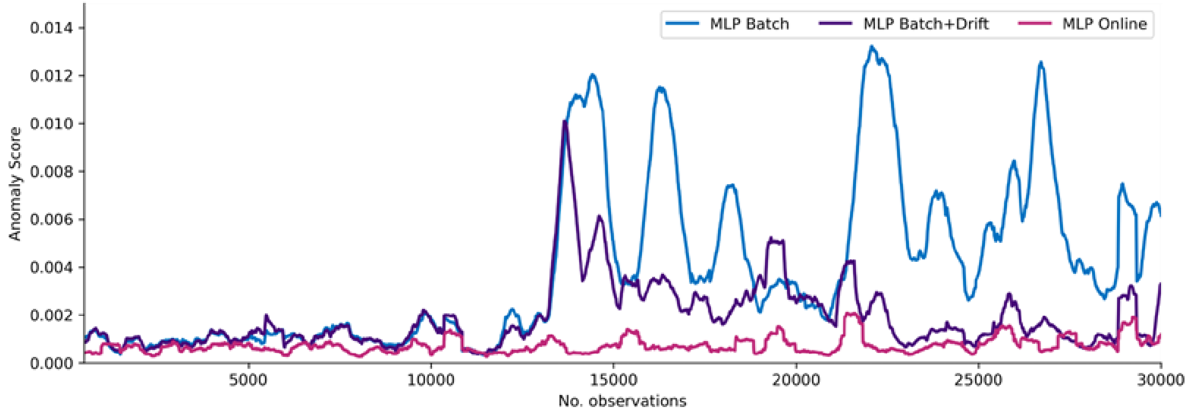
Towards Online Anomaly Detection in Steel Manufacturing Process

We train ML models to predict selected rolling parameters e.g. rolling forces.
The assumption for the outlier detection is that if the measurement differ significantly from model prediction, the anomaly alarm is raised.

Different approaches towards learning process were used to determine optimal learning strategy:

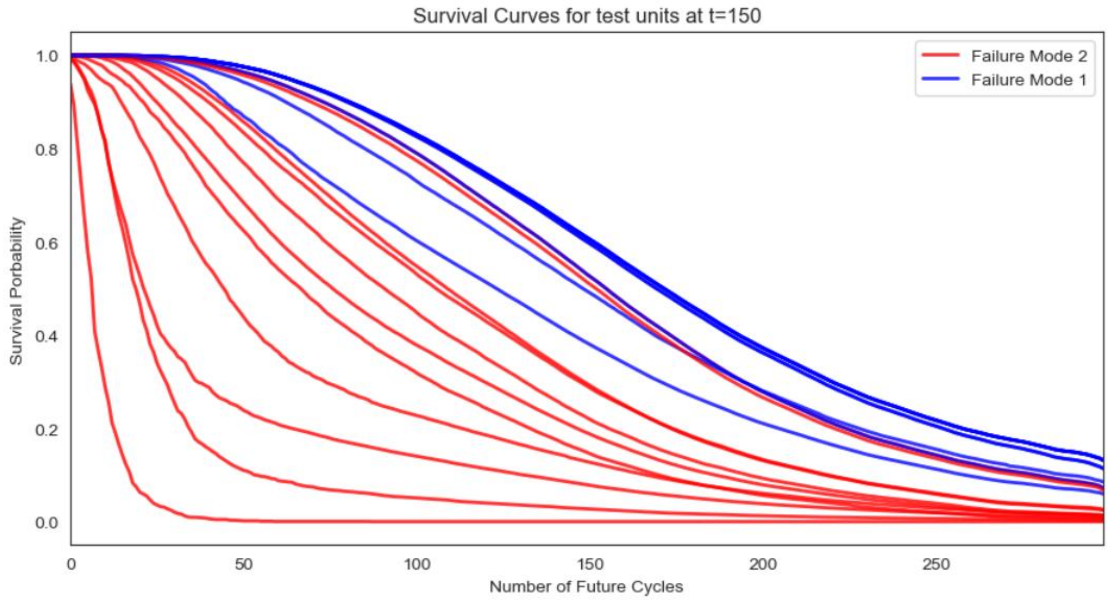
- Batch learning
- Batch learning with concept drift detection
- Online learning

Paper accepted for ICCS 2023 conference.



Understanding Survival Models through Counterfactual Explanations

- Simple survival models e.g., Cox Proportional Hazards (CPH), are inherently interpretable.
- There exist several adaptations of state-of-the-art ML methods for survival analysis e.g., Random Survival Forest or SVM, which require explanations to understand reasons behind the prediction.
- Survival models return curves rather than point estimates, which make them more difficult to interpret.
- Work in cooperation with Halmstad University
- **Submitted to DSAA 2023 PRAXAI**



Understanding Survival Models through Counterfactual Explanations

- We propose a method to generate likely and actionable counterfactual explanations for survival models.
- Actionability is achieved by making selected features immutable, e.g., age.
- Likelihood is achieved by using autoencoder to learn the data representation. The reconstruction error is then included in the loss function.
- Two distinct approaches are presented - one with transforming survival functions into survival scores (regression) and the second with survival patterns discovery (classification)

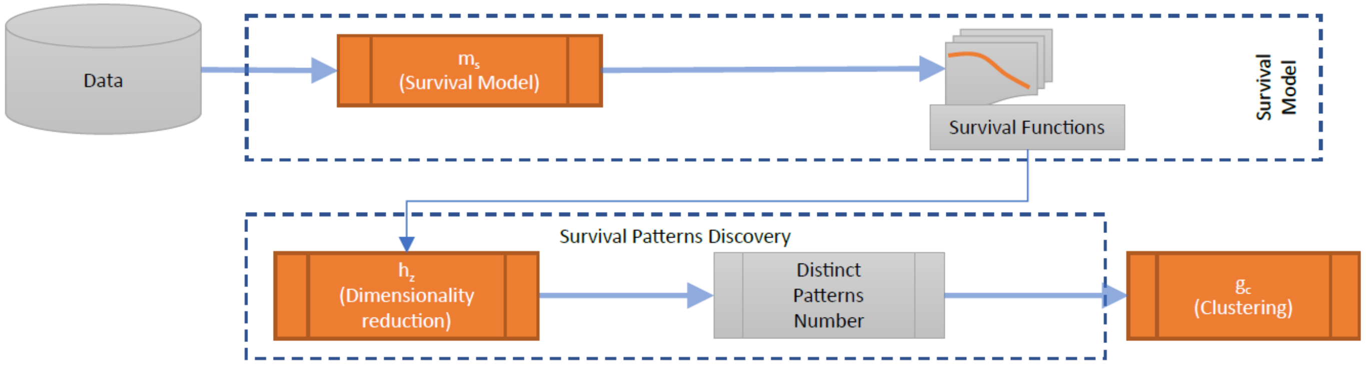
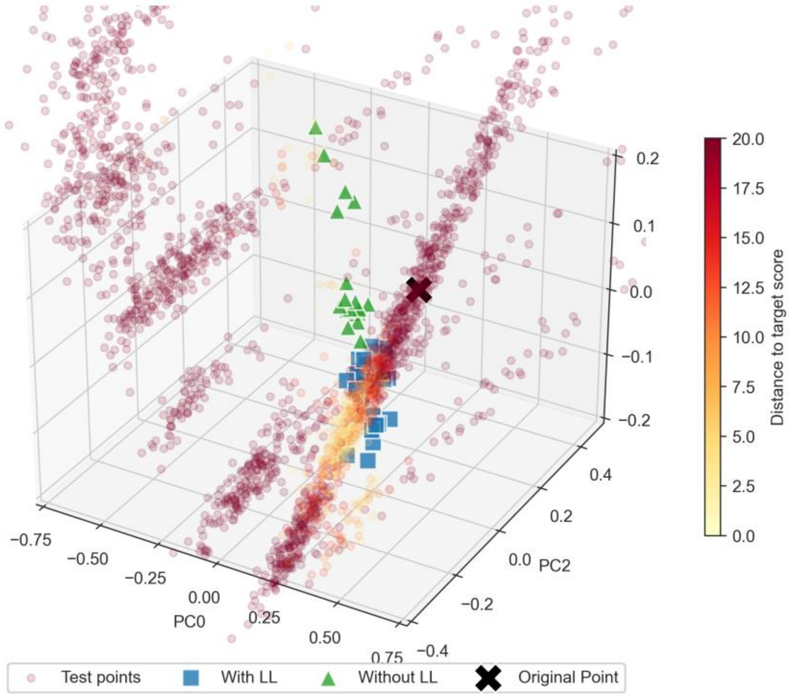
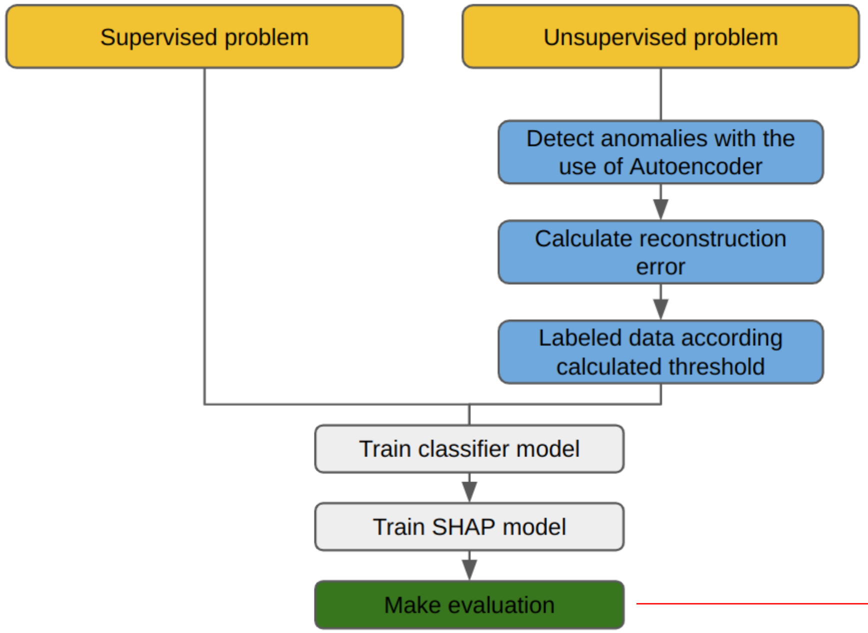
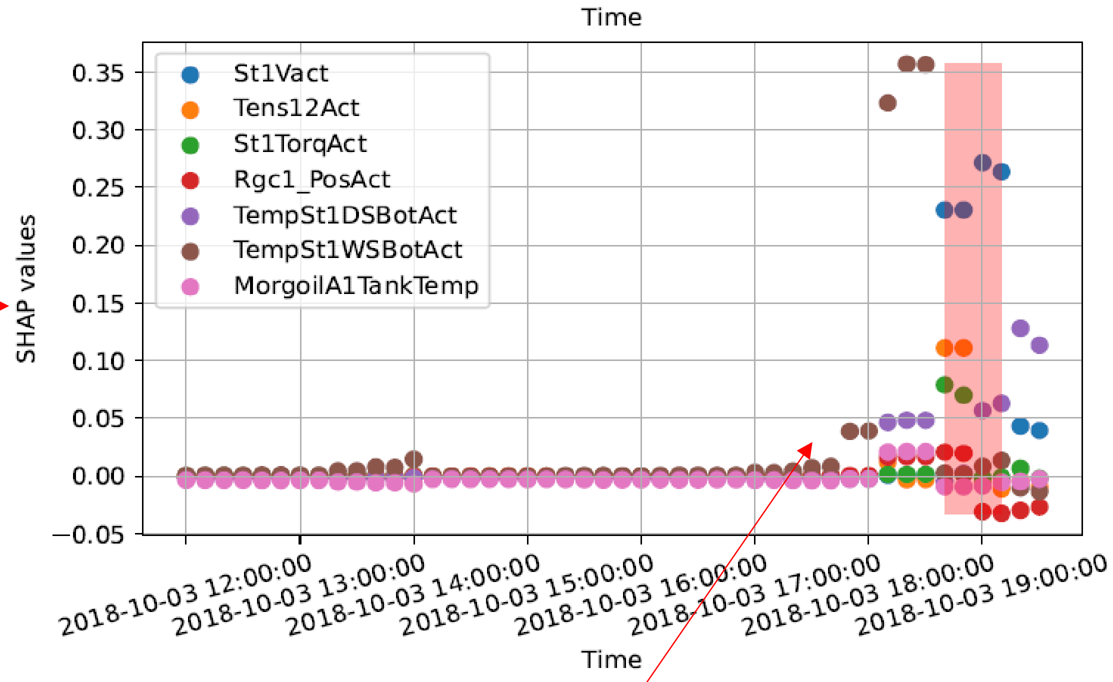


Fig. 1: Patterns discovery Workflow



Feature importance as a tool for root cause analysis in time-series events



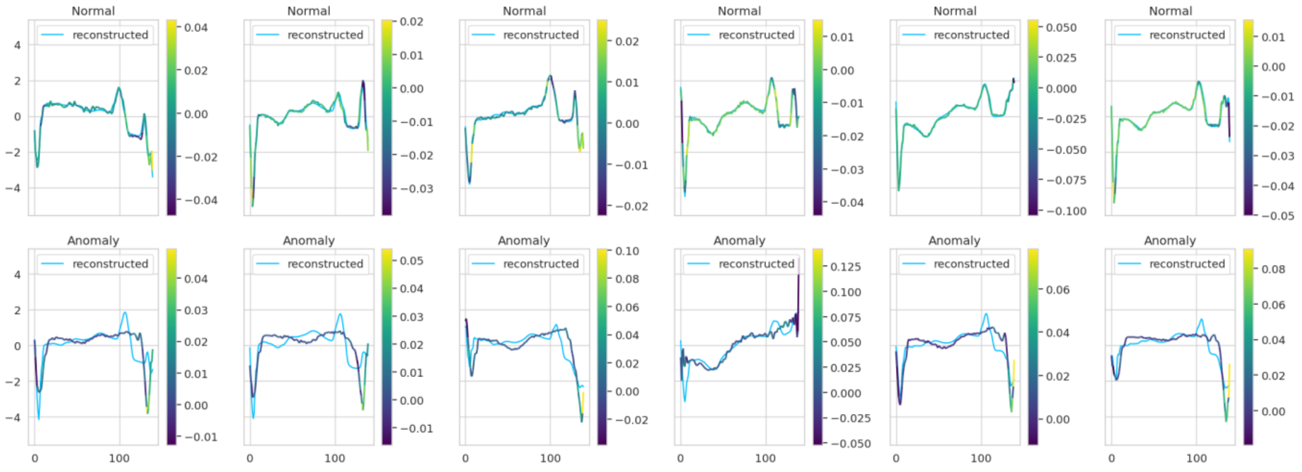


Paper accepted for publication:
Feature importance as a tool for root cause analysis in time-series events
Michał Kuk, Szymon Bobek, Bruno Veloso, Lala Rajaoarisoa and Grzegorz J. Nalepa
INTERNATIONAL CONFERENCE ON COMPUTATIONAL SCIENCE

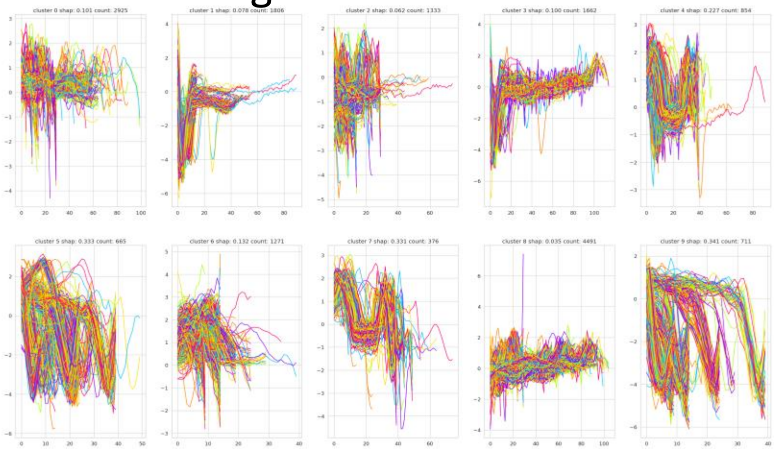
The aim of the work was to demonstrate how the analyses of Shap values near the occurrence of failures can help identify the specific features that led to the failure.

Post-hoc prototype generation and explanation of time series classification

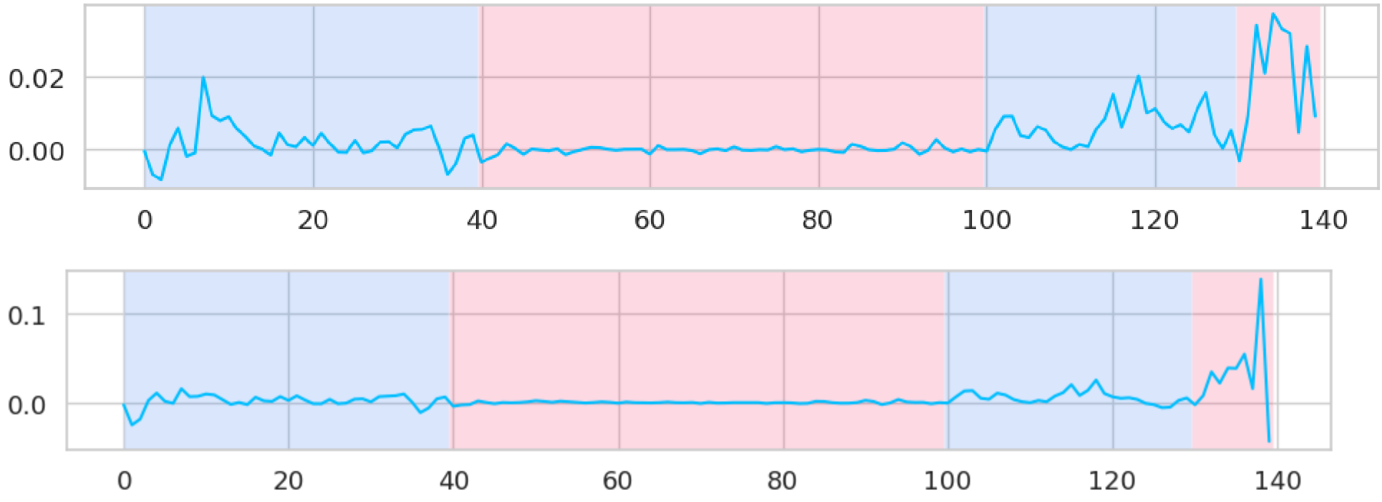
1. Calculate SHAP values



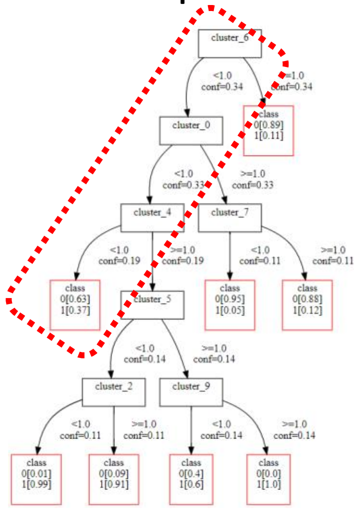
3. Use SHAP segments as mask over real TS & cluster TS segments



2. Segment SHAP values with changepoint detection



3. Build and explainable clssifier based on cluster presence in particualr TS



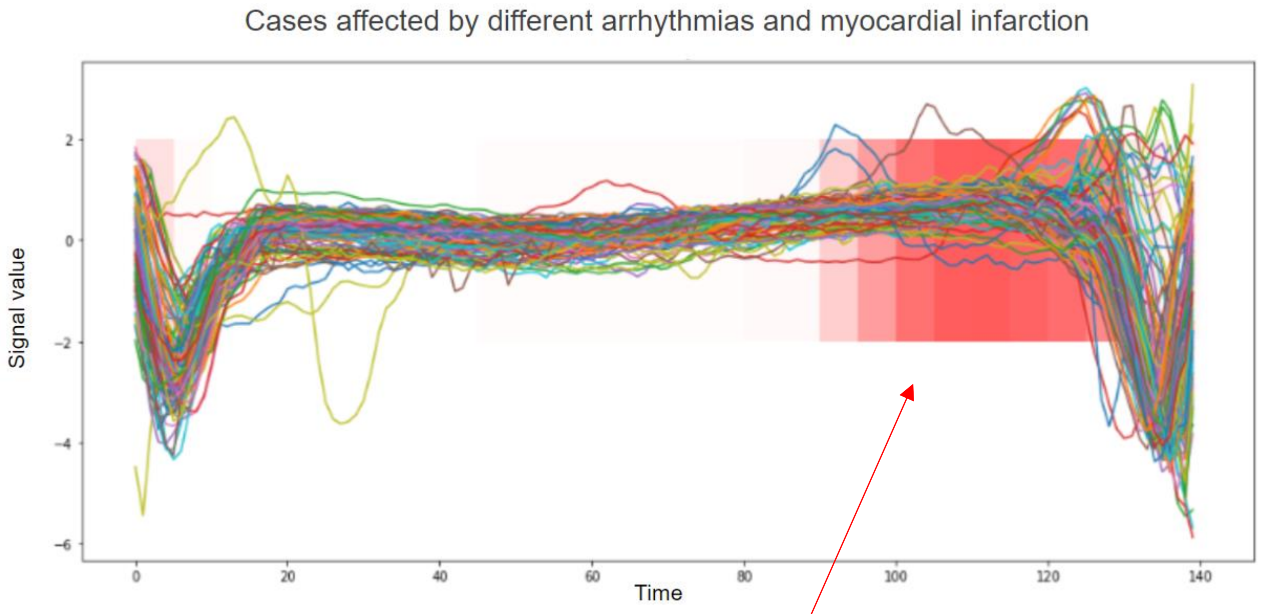
Blackbox model: 0.89 Acc

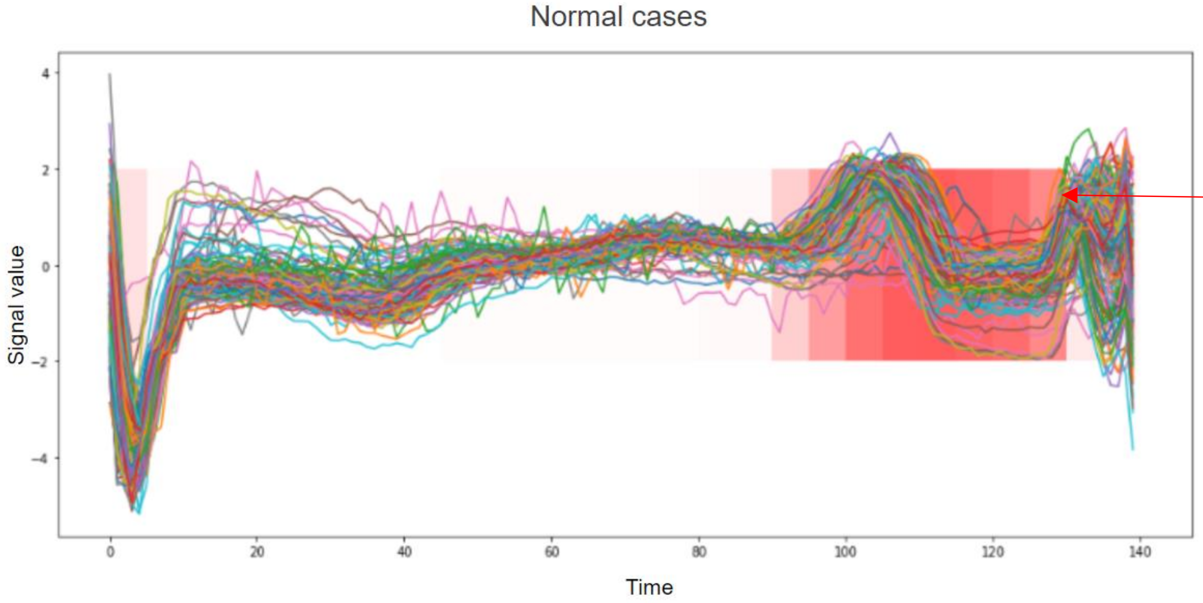
Decision tree: 0.88 Acc

Post-hoc prototype generation and explanation of time series classification

The purpose of the work is to extract segments from a time series, and then apply a change point detection algorithm that explores segments based on shape values and uses this information to explain specific parts of the time series.

The methods should present better rules that should be more understandable to humans, the rules should not be sensitive to small changes in signal values, and the designated prototypes will be able to easily distinguish normal cases from abnormal ones.

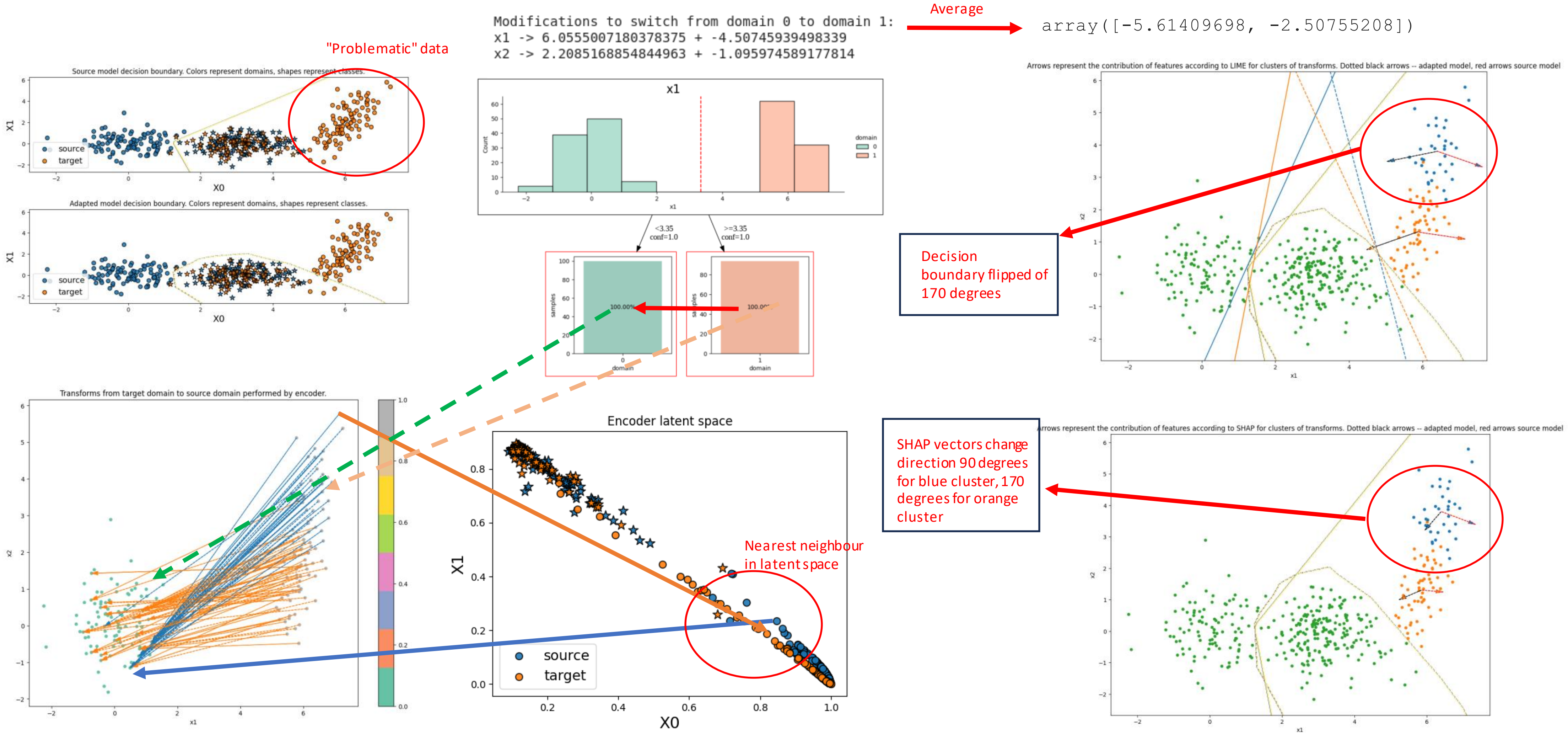




Marked segments which were indicated by the generated rules (LUX algorithm) presented on all analyzed cases.

In this case, the explainable algorithm found the segment which the most differentiates the normal and sick cases. The indicated segments could be treated as a prototype which in a human understanding way presents why the algorithm classifies the signal for normal or not normal ECG.

DAxAI: Explainable Domain Adaptation



Summary

- Knowledge Augmented Clustering (KnAC)
- Local Uncertain eXplanbations (LUX)
- Intelligible XAI (InXAI)
- Technology needs to be human-centric
- Explanations are important for unsupervised methods (KnAC/Explainable clusters)
- *The truth is out there*

Open Challenges in XAI for (not only) Industry 4.0

- Mediating explanations between human and XAI system.
  - Explanation is an act of conveying knowledge
  - Technology needs to be human-centric. Good explanation does not always mean useful or understandable
- Defining mediatable information granules via human-in-the-loop conceptualization.
  - Semantic gap between XAI and different explanation addressee (stakeholders)
- Multi-faced continuous assessment of quality of explanations.
  - Why should I trust... your explanation
  - Correlation does not mean causation

Thank you for your attention!

Give us a feedback @ https://github.com/sbobek/knac

JAGIELLONIAN UNIVERSITY
IN KRAKÓW

https://geist.re

AGH UNIVERSITY OF SCIENCE
AND TECHNOLOGY