The PREFERML project (Proactive Error Avoidance in Production through Machine Learning) was funded by the German Federal Ministry of Education and Research, funding line “Forschung an Fachhochschulen mit Unternehmen (FHProfUnt)“, contract number 13FH249PX6.
We want to thank the company SICK AG for the cooperation and partial funding.

Detection of Concept Drift in Manufacturing Data with SHAP Values to Improve Error Prediction

Authors: Christian Seiffer, Holger Ziekow, Ulf Schreier, Alexander Gerling

Presenter: Christian Seiffer, Furtwangen University

Email: christian.seiffer@hs-furtwangen.de

Person

- 2019, Industrial Engineering (M.Sc.) at Karlsruhe Institute of Technology (KIT)
- Working at a research project „PREFERML“ at Furtwangen University since 2020
- Research areas:

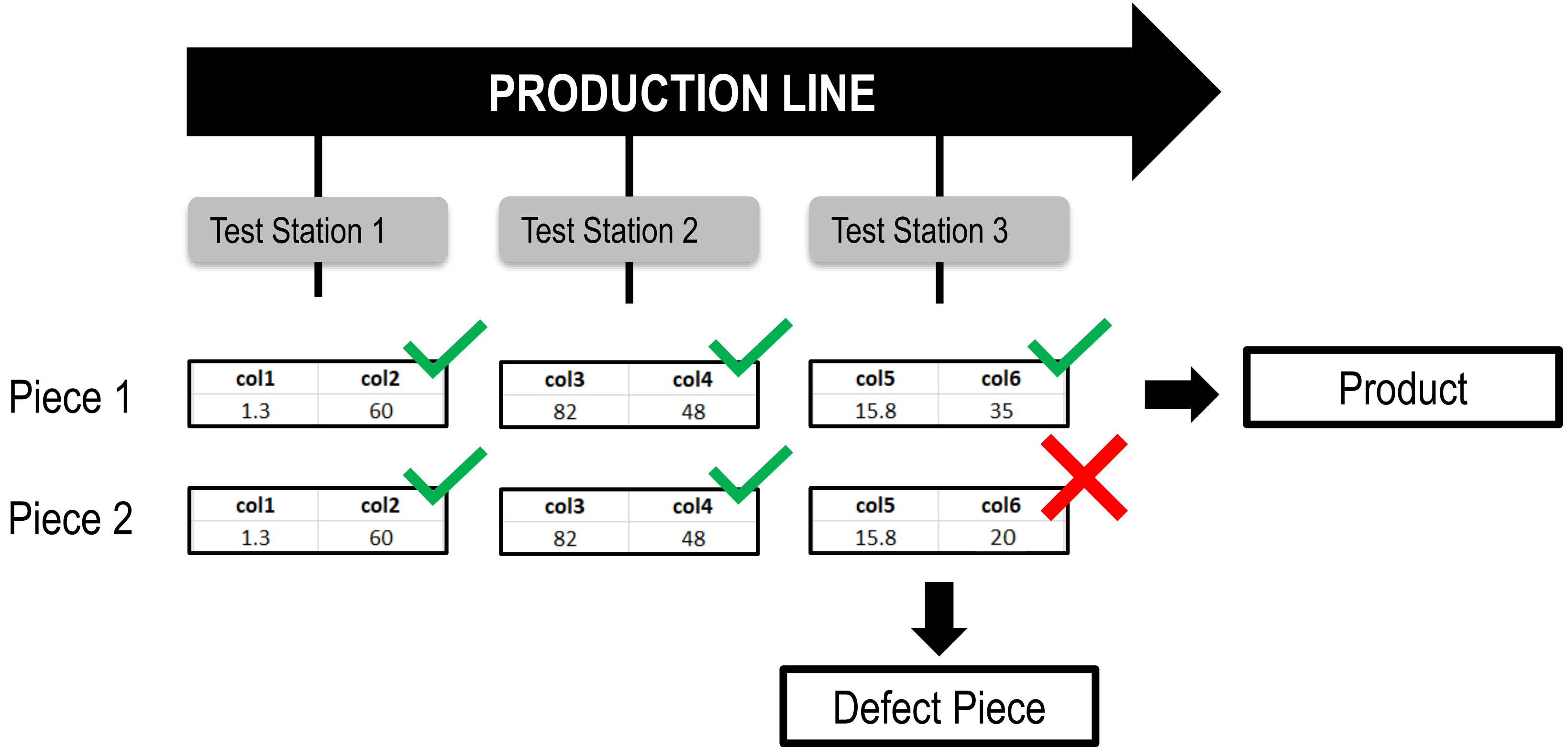
  Error prediction in manufacturing

  Concept Drift in manufacturing

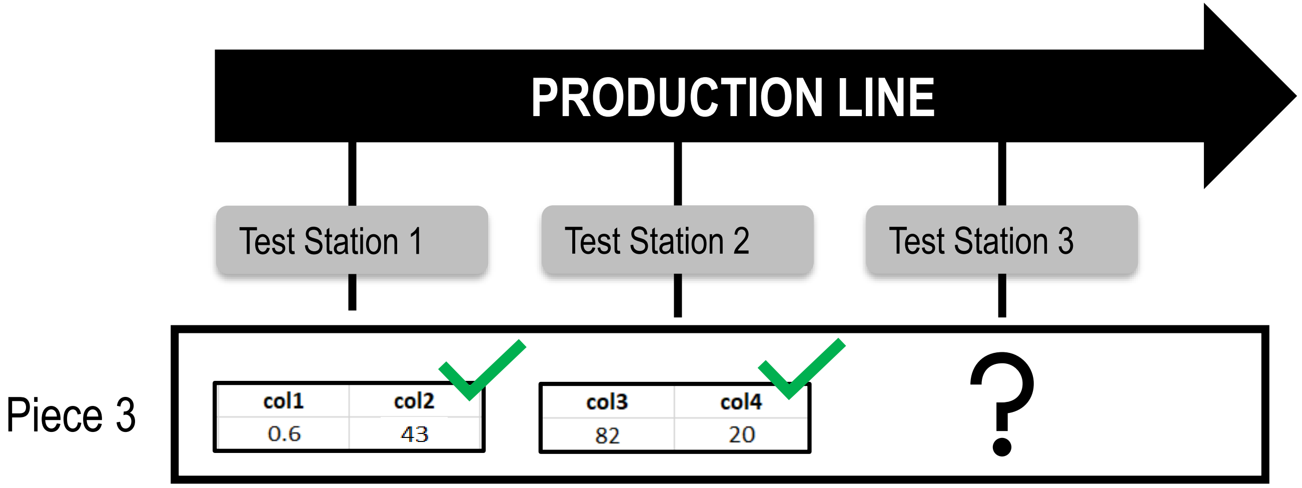
Agenda

- Introduction
- Motivation
- Approach
- Experiments
- Results
- Conclusion

Introduction: Domain



Introduction: Machine Learning in Manufacturing



Machine Learning Classification Model (learned on prior data)

➔ Prediction:

➔ Remove piece at this early stage

- True Error Prediction:
  Avoids errors at an early stage (save costs)
- False Error Prediction:
  Misses profits
- Baseline: No use of ML (i.e., no predictions)

➔ Economic savings depend on TP and FP:

Total savings = TP * c(TP) - FP

➔ Quality Managers: knowledge/guess of c(TP)

➔ Evaluation based on precision and critical threshold τ (corresponding to *Total savings* = 0)

Motivation

- Usually, several cause-error relationships (*Concepts*) in manufacturing
- Production processes subject to dynamic change (e.g. mechanical properties of tools)
  ➔ Cause-error relationships are not permanent ($P_t(Y|X) \neq P_{t+1}(Y|X)$), known as *Concept Drift*
- Hypothesis: *"Adressing concept drift in manufacturing data by separating underlying concepts yields more insights than an overall analysis."*

Goal:

- Detect *Concept Drift* efficiently
- Handle *Concept Drift* individually

Approach – Key Idea

- Cluster data based on corresponding SHAP values[1] – "supervised clustering"
- SHAP value: contribution of a single feature to model output (predicted value)
- Clustering: similar instances with regards to how the model computes predictions for them

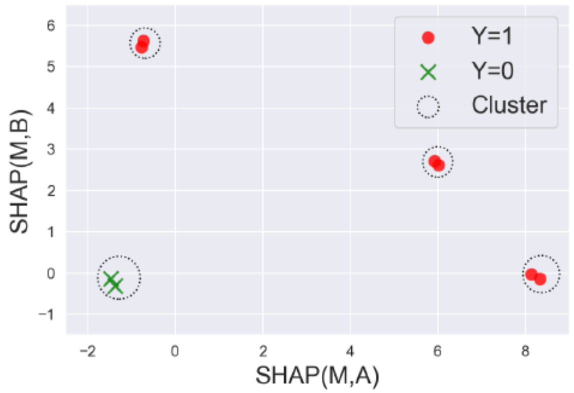
➔ cluster correspond to underlying concepts the model learned

Original Data

ID	A	B
1	0.39	-2.21
2	1.07	0.87
3	-0.63	-0.51
4	-1.47	-0.85
5	0.55	2.61
6	1.16	2.34
7	-0.87	2.43
8	-0.62	2.6

SHAP-transformed data

$SHAP(M,A)$	$SHAP(M,B)$
-1.34	-0.33
-1.28	-0.11
8.36	-0.02
8.33	-0.08
-0.74	5.56
-0.69	5.57
6.00	2.69
6.00	2.69

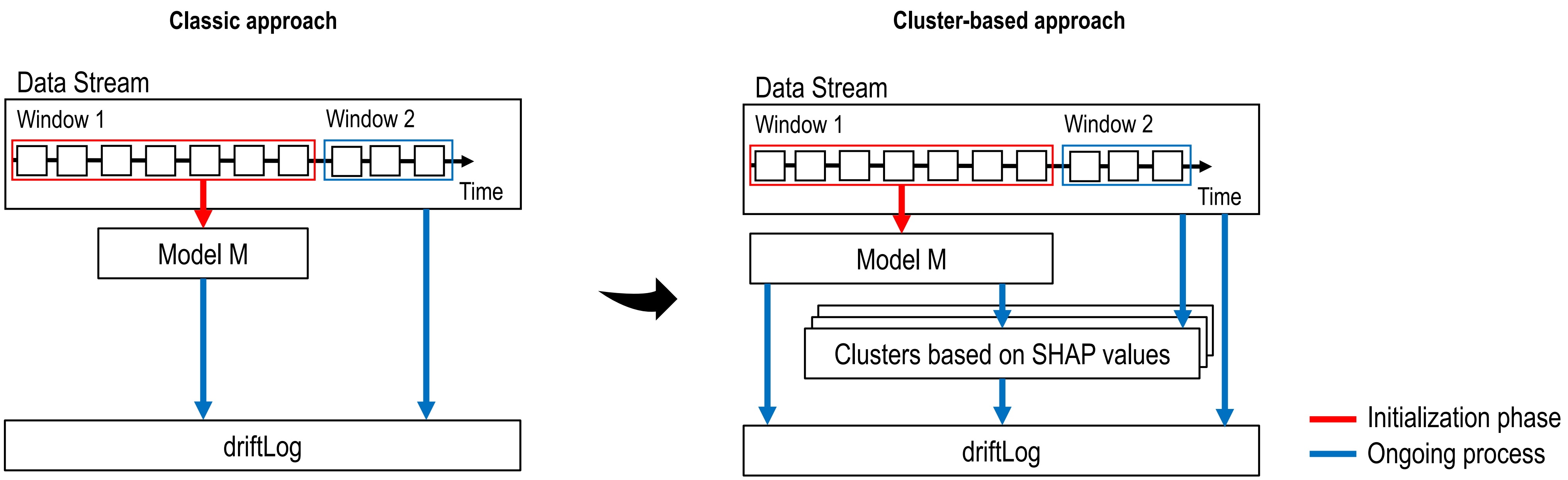


Original Data

ID	A	B	*cluster*
1	0.39	-2.21	1
2	1.07	0.87	1
3	-0.63	-0.51	2
4	-1.47	-0.85	2
5	0.55	2.61	3
6	1.16	2.34	3
7	-0.87	2.43	4
8	-0.62	2.6	4

[1] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Advances in Neural Information Processing Systems, I. Guyon et al., Eds., vol. 30. Curran Associates, Inc., 2017.

Approach – Procedure



- Dividing total data into clusters based on their SHAP values
- driftLog: includes relevant information for drift handling

Approach – Detecting and Handling Drift in Data Streams

- Different clusters yield separate data streams
- Target variable: sliding window of precision over the last n positive predictions $\quad$ *precision* = $\frac{TP}{TP+FP}$
- Decision to be made based on current precision:

  Ignore the next positive prediction or trust it?

Two alternative strategies:

- *handle drift based on current precision only*:

  if precision < critical threshold τ, the next positive prediction is ignored, otherwise not
- *handle drift based on current precision and established drift detectors* (e.g. Page-Hinkley test):

  if precision < τ and drift is detected at the current instance, the following positive predictions are ignored until precision > τ

Experiments

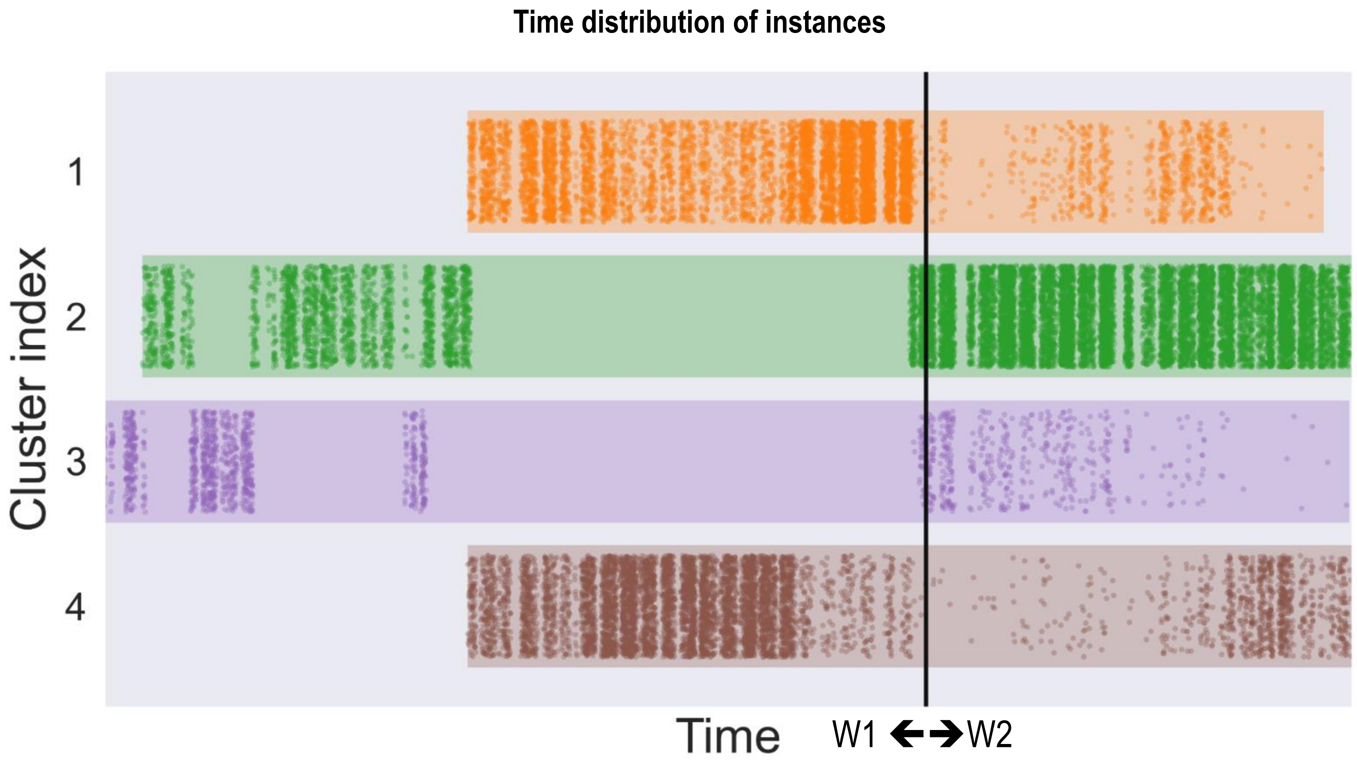
- Real manufacturing data of our industry partner SICK AG
- Data: six production lines, covering about a year
- Each experiment: any type of error at a specific test station, previous features serve as input

  ➔ 30 Experiments, conducted retrospectively, simulation under real-time conditions
- Data containing 67% of errors served as train data, other instances part of window W2
- Classifier: Xgboost with tree booster

  Clustering: k-means (elbow method)

  Drift detector: Page-Hinkley test
- Pretest: suitable parameters (e.g., sliding window size n=100).

Results – Use Case



- Heterogeneous distribution
- Periods where no data is assigned to specific cluster
- Only little data of Cluster 1 in W2
- Data of W2 mostly in Cluster 2

Results – Use Case

Classification results for W2 – basic ML without handling drift

Data	Instances	Errors	TP	FP	TS
Cluster 1	778	16	1	0	10
Cluster 2	11094	301	139	3994	-2604
Cluster 3	479	15	15	463	-313
Cluster 4	993	34	6	0	60
$\sum_{clusters}$	13344	366	161	4457	-2847
Total data	13344	366	161	4457	-2847

- Bad performance in Cluster 2 and Cluster 3
- Cluster 2 affects overall performance due to its size

Handle drift based on current precision only

Data	Instances	Errors	TP	FP	TS
Cluster 1	778	16	1	0	10
Cluster 2	11094	301	22	195	25
Cluster 3	479	15	9	91	-1
Cluster 4	993	34	6	0	60
$\sum_{clusters}$	13344	366	38	286	94
Total Data	13344	366	31	307	3

🡸**Best result**

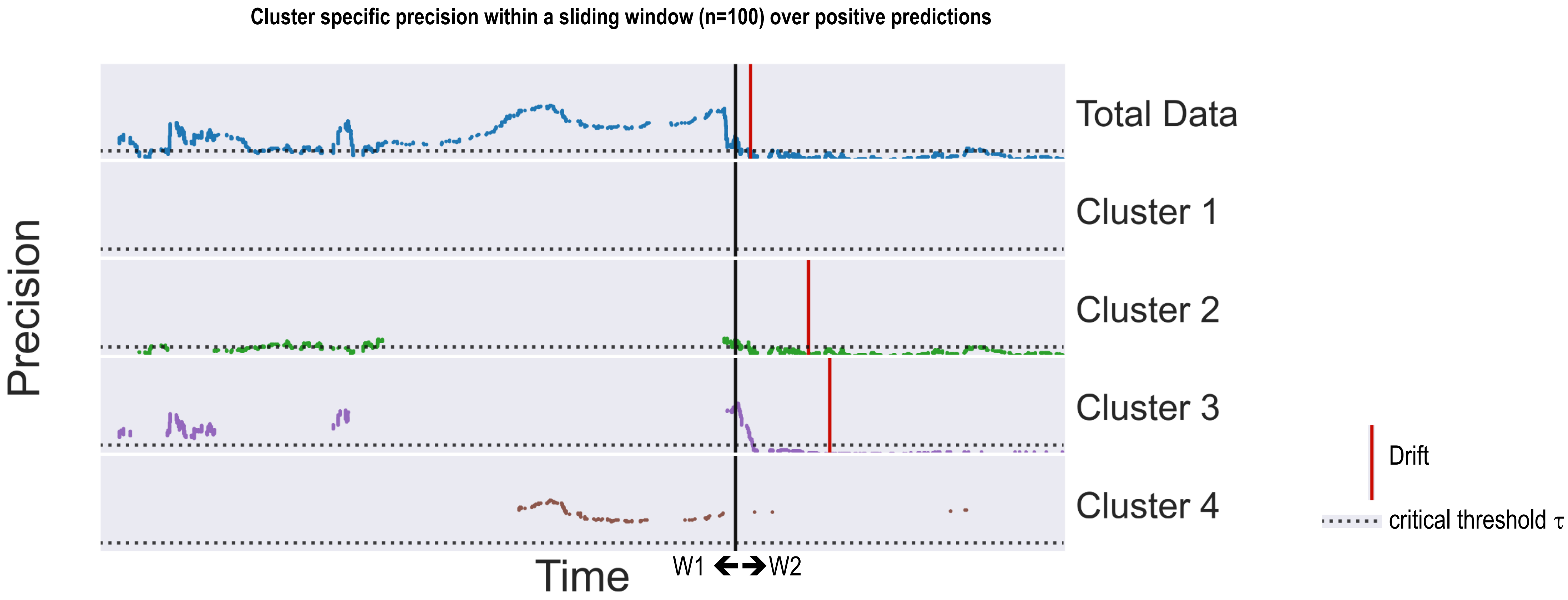
- Performance of Cluster 2 turned into benefit
- Cluster-specific approach outperforms classical approach

Handle drift based on current precision and drift detector

Data	Instances	Errors	TP	FP	TS
Cluster 1	778	16	1	0	10
Cluster 2	11094	301	29	272	18
Cluster 3	479	15	10	99	1
Cluster 4	993	34	6	0	60
$\sum_{clusters}$	13344	366	46	372	89
Total Data	13344	366	47	560	-113

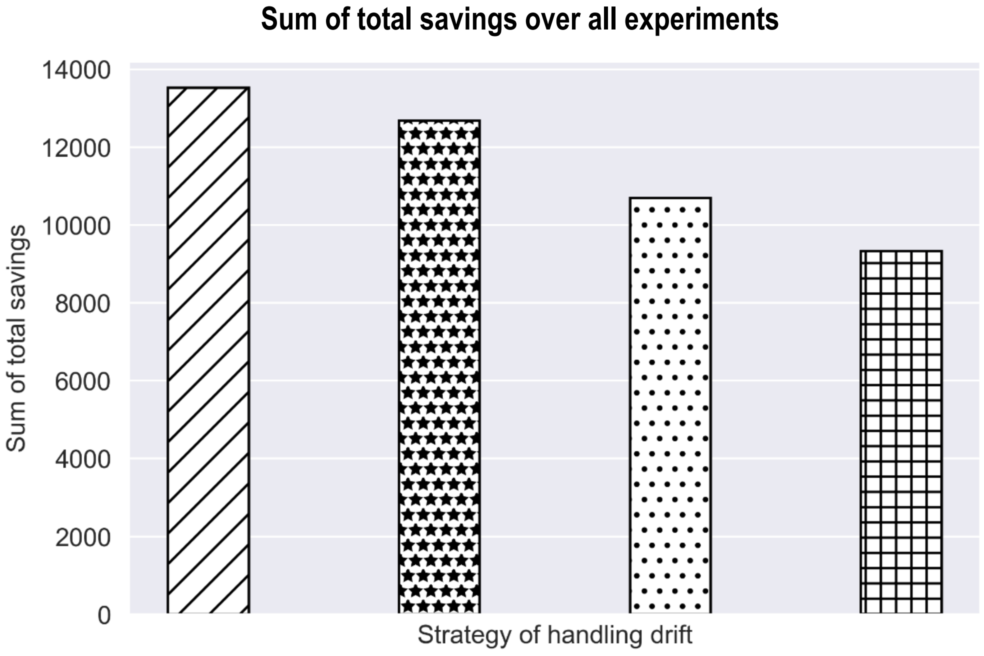
- Performance of all clusters turned into benefit
- Cluster-specific approach outperforms classical approach

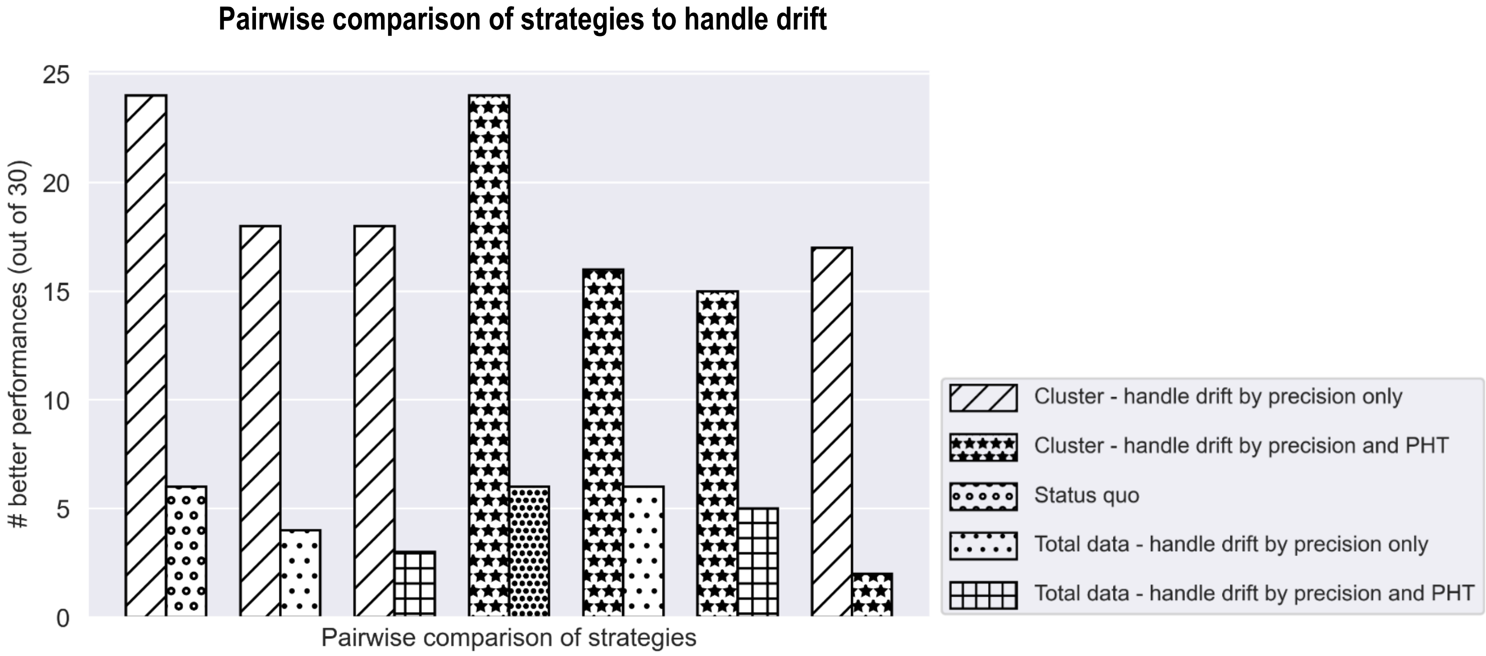
Results – Use Case



- (Too) little data for Cluster 1, good performance of Cluster 4
- Cluster 3: deterioration at the beginning of W2
- Usually no lasting improvement after deterioration below τ

Results – General View





- Cluster-specific approach outperforms classical approach
- Handle drift by precision only seems more promising

Conclusion

- Method that uses SHAP values to assign the learned concepts to clusters so that they can be examined individually
- Cluster specific assessment in combination with two strategies for handling drift
- Clustering based assessment outperformed approaches without clustering
  (for both strategies of handling drift)