UNIVERSITÄT BIELEFELD
Technische Fakultät

Machine Learning in the context of drift

Barbara Hammer

Machine Learning Group, Bielefeld University

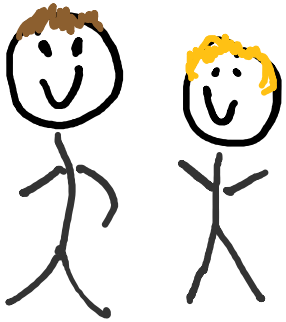
Colloquium Talk at Potsdam University, 2022

UNIVERSITÄT BIELEFELD
Technische Fakultät

Learning without drift

UNIVERSITÄT BIELEFELD
Technische Fakultät

Learning without drift

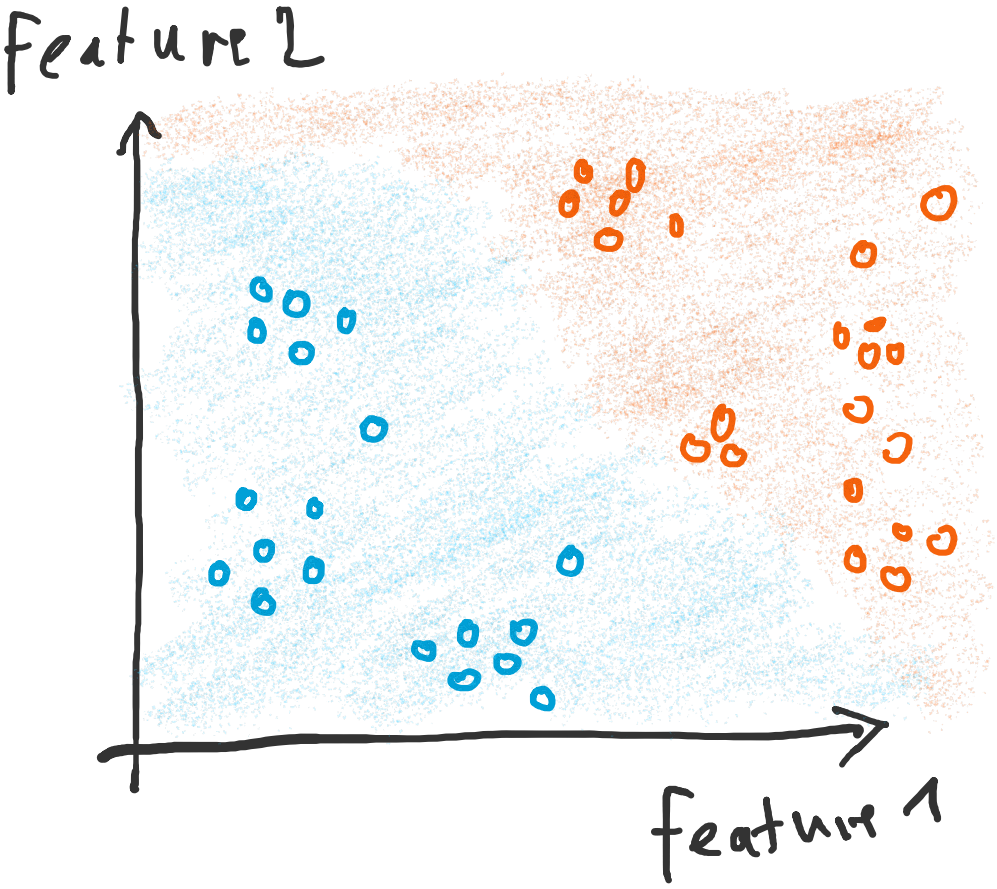


UNIVERSITÄT BIELEFELD
Technische Fakultät

Learning without drift

UNIVERSITÄT BIELEFELD
Technische Fakultät

Learning without drift



Given **training data**

$D = \{(x^1, y^1), \ldots, (x^n, y^n)\} \subset X \times Y$

sampled i.i.d. w.r.t. probability measure $P = P_{X \times Y}$ where $X = \mathbb{R}^d$, $Y = \{1,2\}$,

we aim for a **classification prescription**

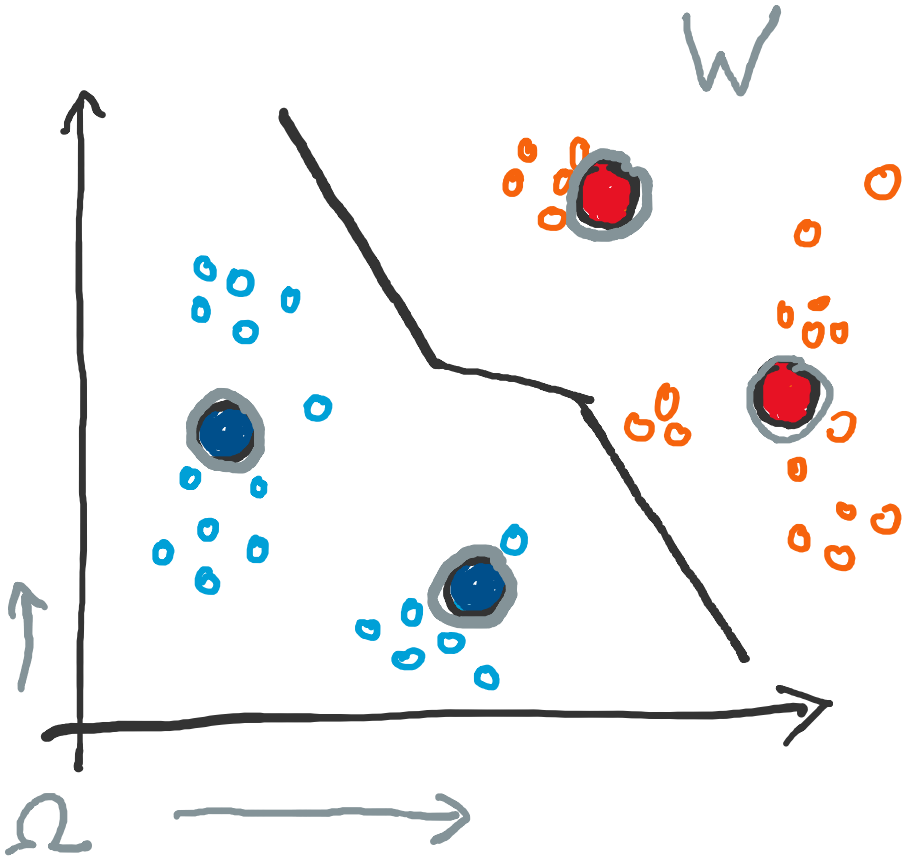
$f_\theta : X \to Y$

minimizing $E_P\big(l(f_\theta(x), y)\big)$ with **0-1-loss**

$$l: Y^2 \to \mathbb{R},\ (y', y) \mapsto 1 - \delta_{y',y} = \begin{cases} 0 & y' = y \\ 1 & y' \neq y \end{cases}$$

UNIVERSITÄT BIELEFELD
Technische Fakultät

GMLVQ - function



GMLVQ (generalized matrix learning vector quantization) – **function**:

Prototypes $(w^1, l^1), \ldots, (w^p, l^p) \in X \times Y$ and **weight matrix** $\Omega \in \mathbb{R}^{d \times d}$ induce a **winner-takes-it-all classification**

$f_{W,\Omega} : x \mapsto l^{I(x)}$ with

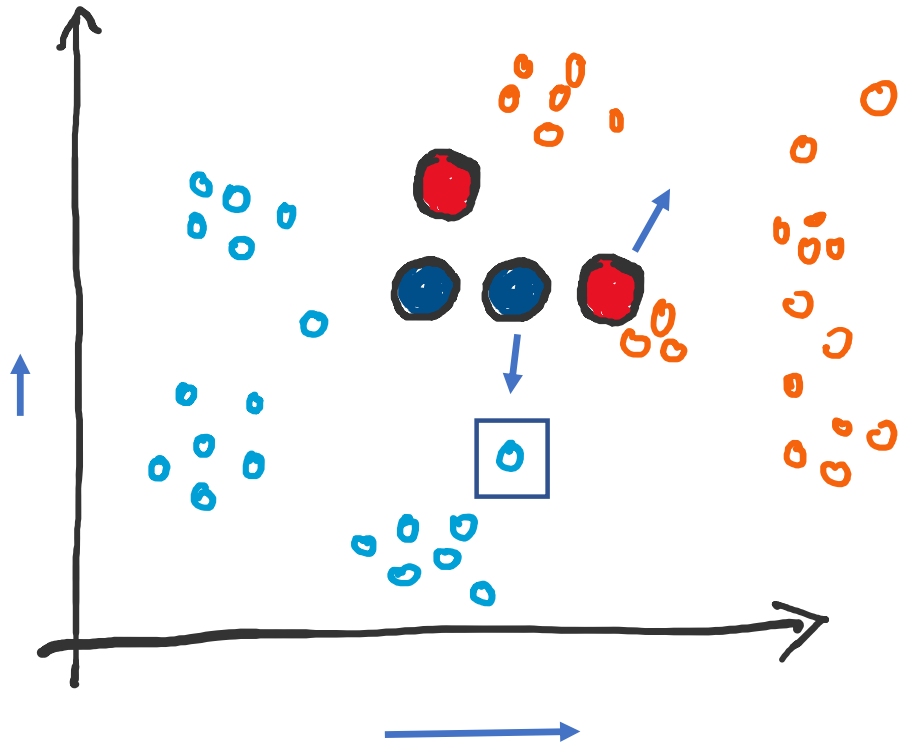
winner $I(x) = \mathrm{argmin}_i \, d_\Omega(x, w^i)$ and

distance $d_\Omega(x, w) = |\Omega(x - w)|^2$

Petra Schneider, Michael Biehl, Barbara Hammer: Adaptive Relevance Matrices in Learning Vector Quantization. Neural Computation 21(12): 3532-3561 (2009), code available at https://github.com/si-cim/prototorch, https://github.com/si-cim/protoflow

UNIVERSITÄT BIELEFELD
Technische Fakultät

GMLVQ - training



GMLVQ – training:

Given a training set, optimize an **approximation of the empirical loss** w.r.t. parameters W, Ω with $\det \Omega = 1$

$$E_{GMLVQ} := \sum_i \Phi\left(\frac{d_\Omega(x^i,w^+)-d_\Omega(x^i,w^-)}{d_\Omega(x^i,w^+)+d_\Omega(x^i,w^-)}\right)$$

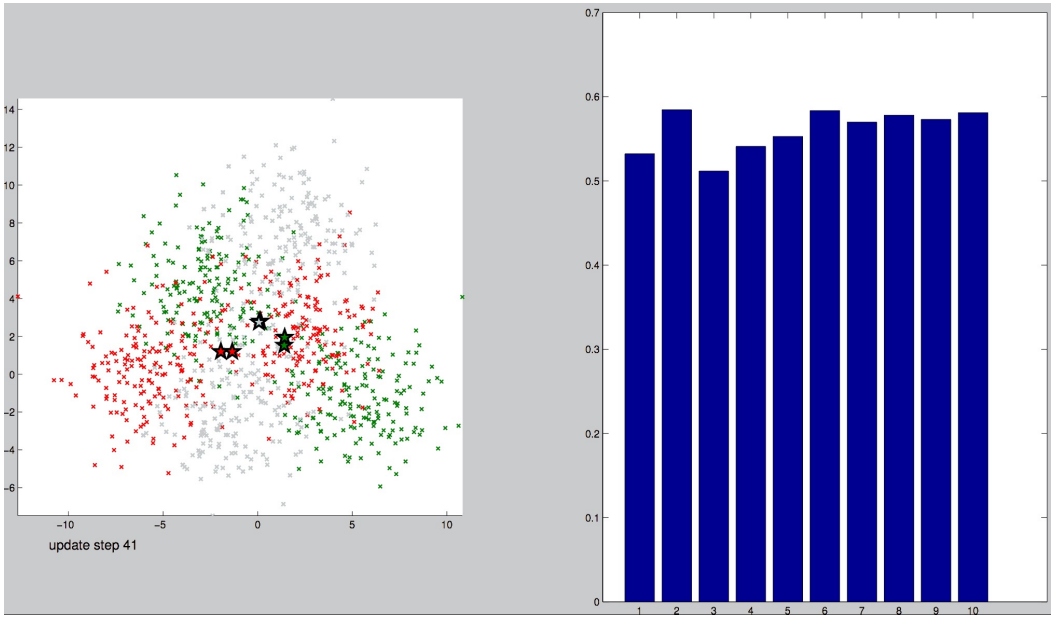
Φ is a monotonic function, $w^{\pm}$ is the closest prototype with correct / incorrect label, and the **sample margin** is

$$M(x^i) := d_\Omega(x^i, w^-) - d_\Omega(x^i, w^+)$$

Petra Schneider, Michael Biehl, Barbara Hammer: Adaptive Relevance Matrices in Learning Vector Quantization. Neural Computation 21(12): 3532-3561 (2009), code available at https://github.com/si-cim/prototorch, https://github.com/si-cim/protoflow

UNIVERSITÄT BIELEFELD
Technische Fakultät

GMLVQ - training



GMLVQ – training:

Given a training set, optimize an **approximation of the empirical loss** w.r.t. parameters W, Ω with $\det \Omega = 1$

$$E_{GMLVQ} := \sum_i \Phi\left(\frac{d_\Omega(x^i,w^+)-d_\Omega(x^i,w^-)}{d_\Omega(x^i,w^+)+d_\Omega(x^i,w^-)}\right)$$

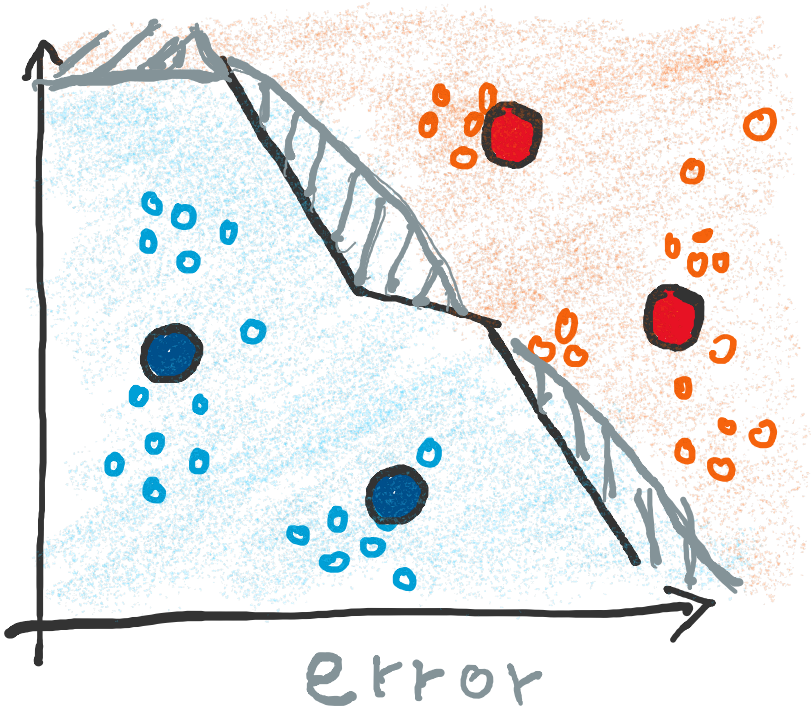
Φ is a monotonic function, $w^{\pm}$ is the closest prototype with correct / incorrect label, and the **sample margin** is

$$M(x^i) := d_\Omega(x^i, w^-) - d_\Omega(x^i, w^+)$$

Petra Schneider, Michael Biehl, Barbara Hammer: Adaptive Relevance Matrices in Learning Vector Quantization. Neural Computation 21(12): 3532-3561 (2009), code available at https://github.com/si-cim/prototorch, https://github.com/si-cim/protoflow

UNIVERSITÄT BIELEFELD
Technische Fakultät

GMLVQ - guarantees



GMLVQ – generalization ability:

Assume data is bounded by B, consider the margin loss function with parameter ρ

$$L_\rho : \mathbb{R} \to \mathbb{R}, t \mapsto \begin{cases} 1 & t \leq 0 \\ 1 - t/\rho & 0 \leq t \leq \rho \\ 0 & t \geq \rho \end{cases}$$

and the **empirical margin loss** $\widehat{E_n^\rho} := \sum_i L_\rho(M(x^i))/n$

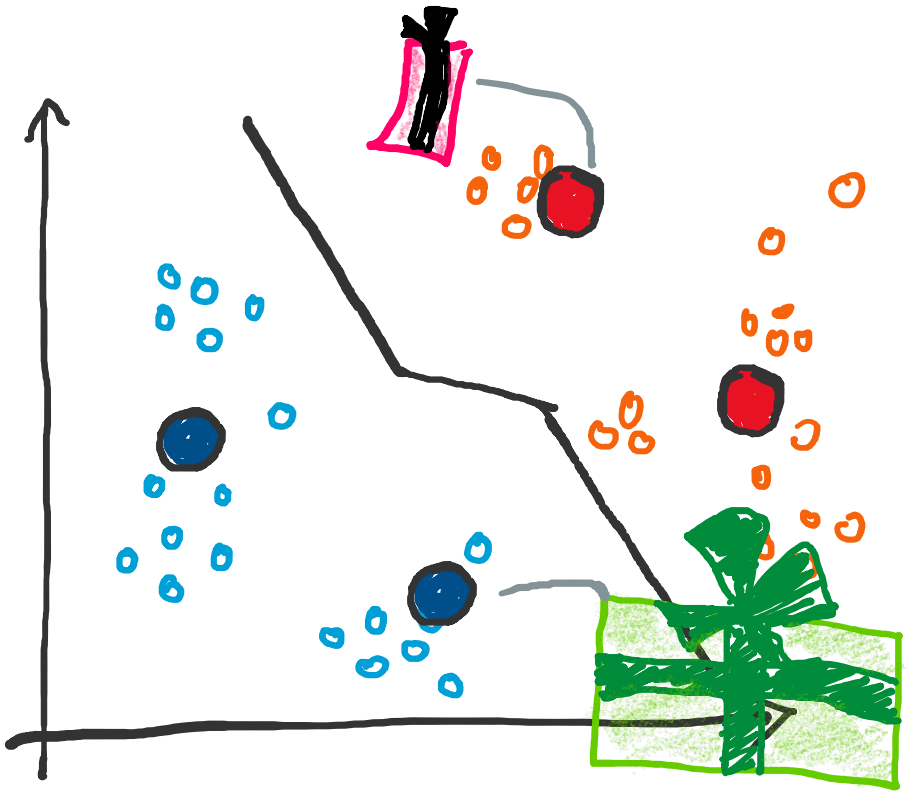
With probability δ the expected loss is bounded by

$$E_P\left(l\left(f_{W,\Omega}(x), y\right)\right) \leq \widehat{E_n^\rho} + \frac{1}{\sqrt{n}} \mathcal{O}\left(\frac{p^{\frac{3}{2}} B^3}{\rho} + \frac{\sqrt{\ln(\frac{1}{\delta})}}{\min\{1,\rho\}}\right)$$

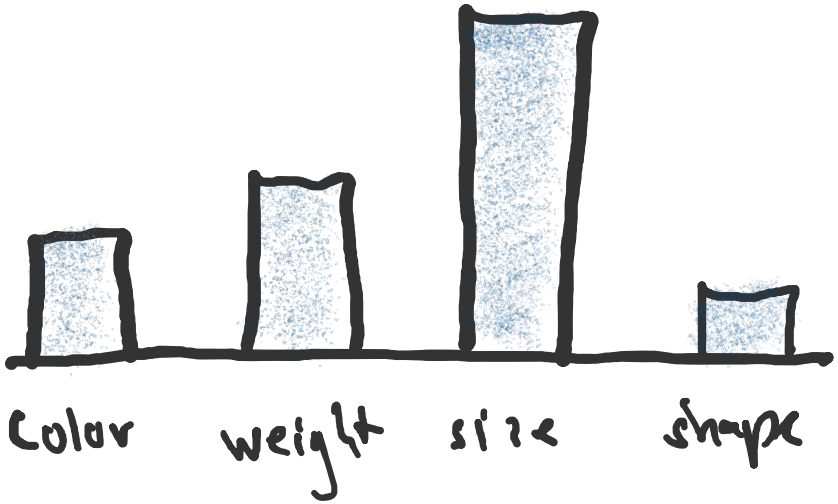
Petra Schneider, Michael Biehl, Barbara Hammer: Adaptive Relevance Matrices in Learning Vector Quantization. Neural Computation 21(12): 3532-3561 (2009)

UNIVERSITÄT BIELEFELD
Technische Fakultät

GMLVQ - interpretability



GMLVQ constitutes a natively **interpretable model** by inspecting prototypes W and relevance terms $\mathrm{diag}(\Omega^T\Omega)$.

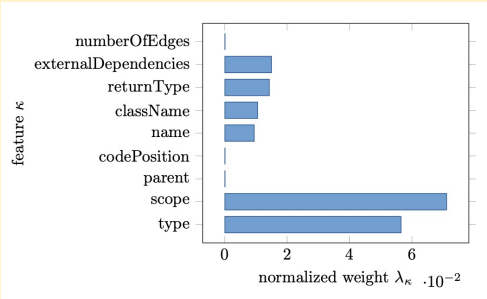


e.g. Arlt/Stewart, PCT/GB2010/000274: Assay for detection of adrenal tumor

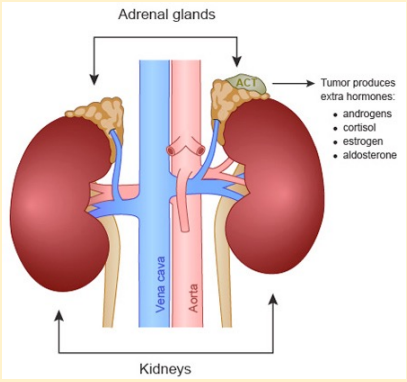
UNIVERSITÄT BIELEFELD
Technische Fakultät

GMLVQ - applications

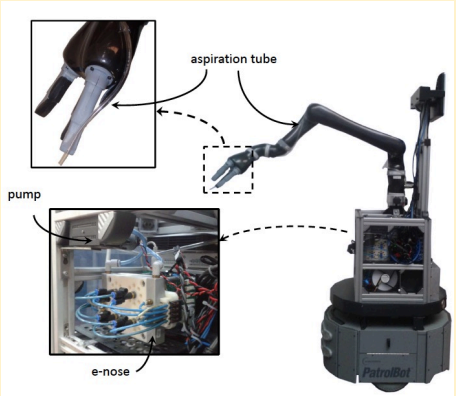
Viktor Losing, Barbara Hammer, Heiko Wersing: Interactive online learning for obstacle classification on a mobile robot. IJCNN 2015: 1-8



Benjamin Paaßen, Bassam Mokbel, Barbara Hammer: Adaptive structure metrics for automated feedback provision in intelligent tutoring systems. Neurocomputing 192: 3-13 (2016)



Lydia Fischer, Barbara Hammer, Heiko Wersing: Optimal local rejection for classifiers. Neurocomputing 214: 445-457 (2016)



Johannes Brinkrolf, Barbara Hammer: Time integration and reject options for probabilistic output of pairwise LVQ. Neural Comput. Appl. 32(24): 18009-18022 (2020)

UNIVERSITÄT BIELEFELD
Technische Fakultät

Learning with drift

- Transfer learning
- Continuous adaptation
- A few thoughts on drift

UNIVERSITÄT BIELEFELD
Technische Fakultät

Transfer learning I

UNIVERSITÄT BIELEFELD
Technische Fakultät

SEMG control of prosthesis

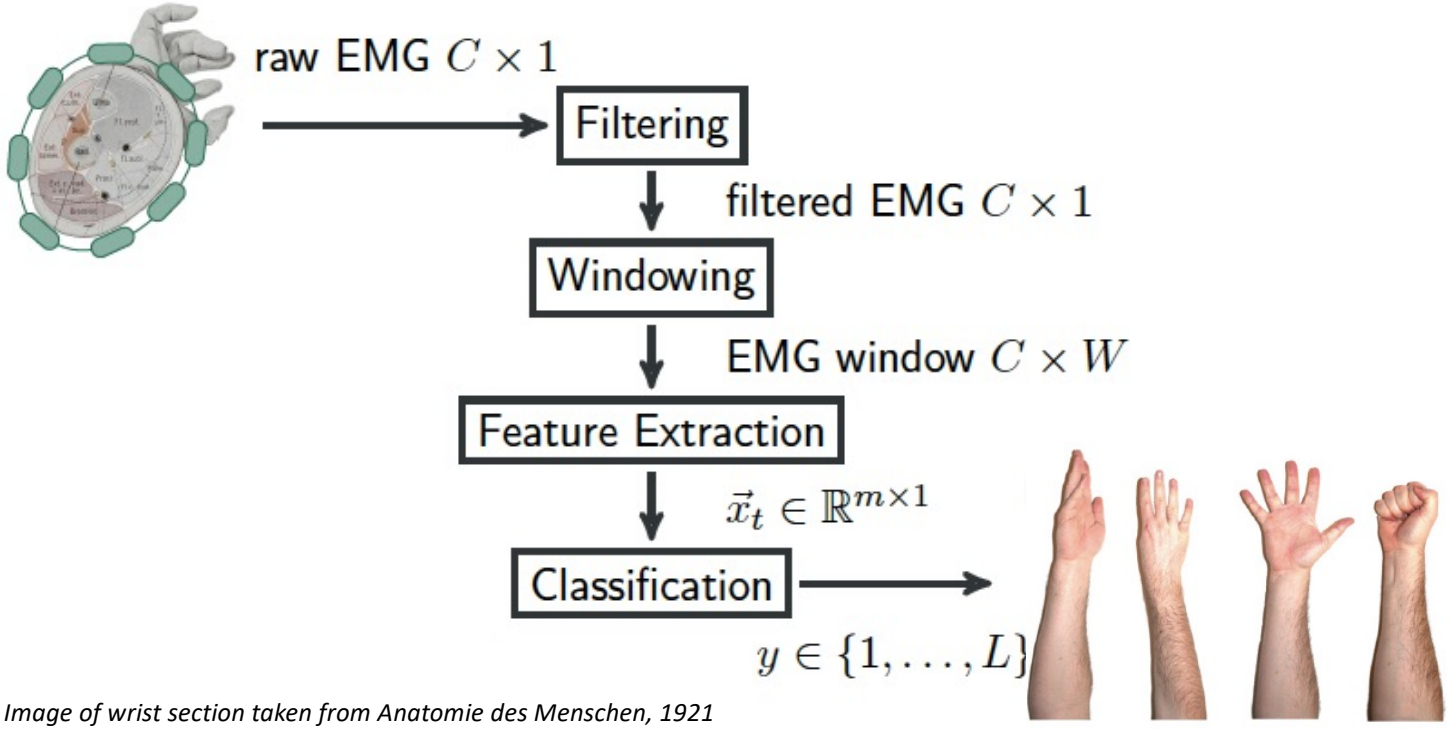


Image of wrist section taken from Anatomie des Menschen, 1921
Image of hand taken from Fluctuating EMG signals: Investigating long-term effects of pattern matching algorithms P Kaufmann, K Englehart, M Platzner - 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology , 2010

UNIVERSITÄT BIELEFELD
Technische Fakultät

Classification of sEMG data

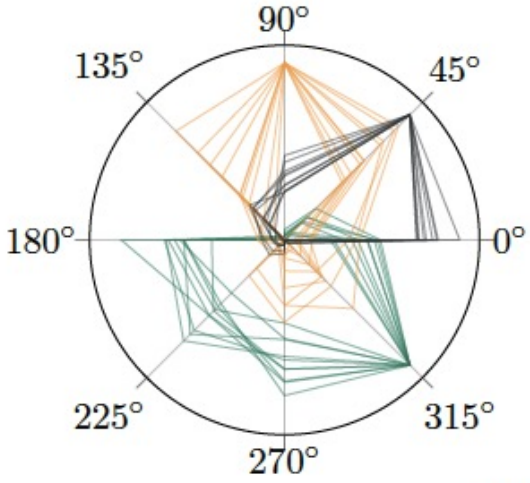
pronation/supination	0.6% (0.3%)
flexion/extension	0.5% (0.5%)
open/close	0.7% (0.4%)

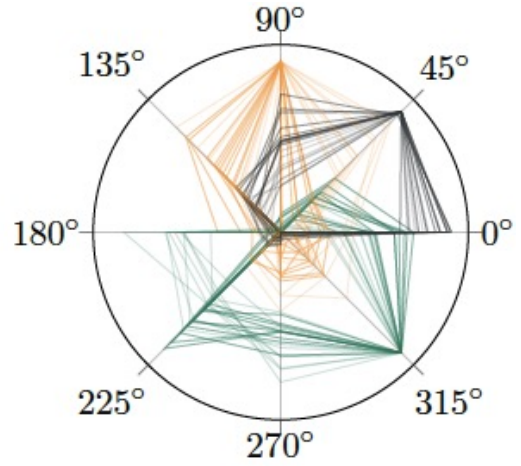
Benjamin Paaßen, Alexander Schulz, Janne Hahne, Barbara Hammer: Expectation maximization transfer learning and its application for bionic hand prostheses. Neurocomputing 298: 122-133 (2018)

UNIVERSITÄT BIELEFELD
Technische Fakultät

Shift of sensors

Image of wrist section taken from Anatomie des Menschen, 1921





UNIVERSITÄT BIELEFELD
Technische Fakultät

Shift of sensors

- Effect of 8mm transversal shift

degree of freedom	original	shift
pronation/supination	0.6% (0.3%)	36.2% (15.4%)
flexion/extension	0.5% (0.5%)	10.9% (3.7%)
open/close	0.7% (0.4%)	29.1% (6.4%)

Benjamin Paaßen, Alexander Schulz, Janne Hahne, Barbara Hammer: Expectation maximization transfer learning and its application for bionic hand prostheses. Neurocomputing 298: 122-133 (2018)

UNIVERSITÄT BIELEFELD
Technische Fakultät

Supervised transfer learning

Transfer learning:

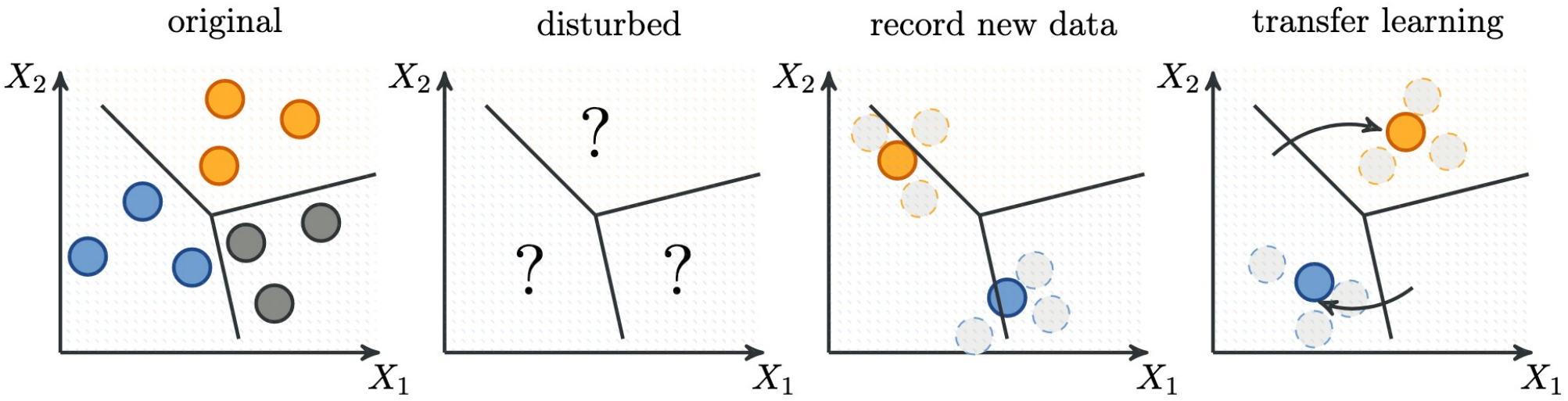
Given a function $f: X \rightarrow Y$

Given (few) samples $(x^i, g(x^i))$ where $g: X' \rightarrow Y$ is different from f but shares some structure

Learn g from the given data with the help of f

UNIVERSITÄT BIELEFELD
Technische Fakultät

Metric-based transfer learning



UNIVERSITÄT BIELEFELD
Technische Fakultät

Metric-based transfer learning

LVQ Model:

$$\boldsymbol{W} \leftarrow \underset{\boldsymbol{W}}{\operatorname{argmin}} E_{\mathrm{GMLVQ}}(\boldsymbol{W}, \boldsymbol{X})$$

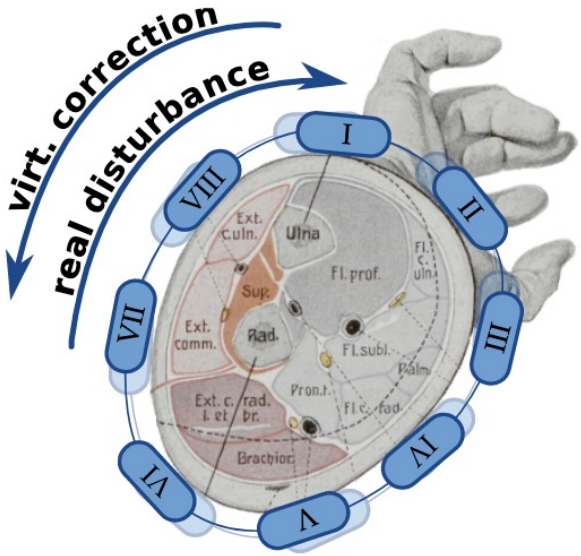
Transfer Learning:

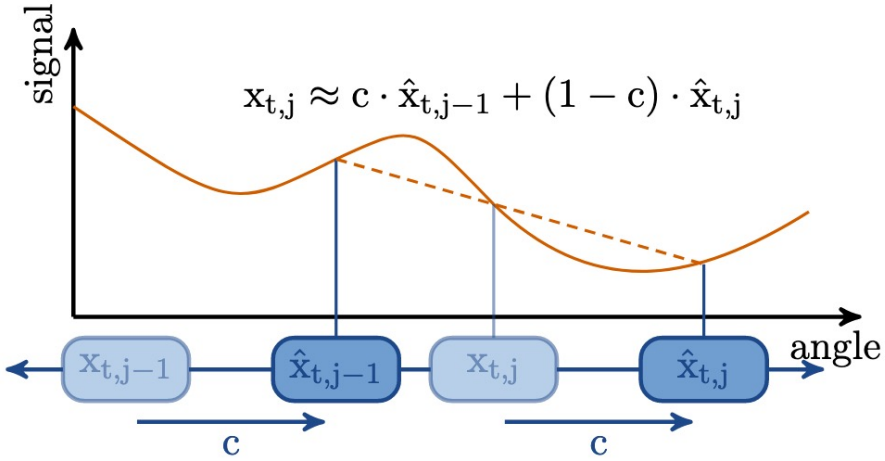
$$\boldsymbol{H}_{\mathsf{p}} \leftarrow \underset{\boldsymbol{H}_{\mathsf{p}}}{\operatorname{argmin}} E(\boldsymbol{W}, \boldsymbol{H}_{\mathsf{p}} \cdot \hat{\boldsymbol{X}})$$

Counteracting Electrode Shifts in Upper-Limb Prosthesis Control via Transfer Learning, *Prahm C, Schulz A, Paaßen B, Schoisswohl J, Kaniusas E, Dorffner G, Hammer B, Aszmann O (2019) IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27(5): 956-962.

UNIVERSITÄT BIELEFELD
Technische Fakultät

Metric-based transfer learning





Counteracting Electrode Shifts in Upper-Limb Prosthesis Control via Transfer Learning, *Prahm C, Schulz A, Paaßen B, Schoisswohl J, Kaniusas E, Dorffner G, Hammer B, Aszmann O (2019) IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27(5): 956-962.

UNIVERSITÄT BIELEFELD
Technische Fakultät

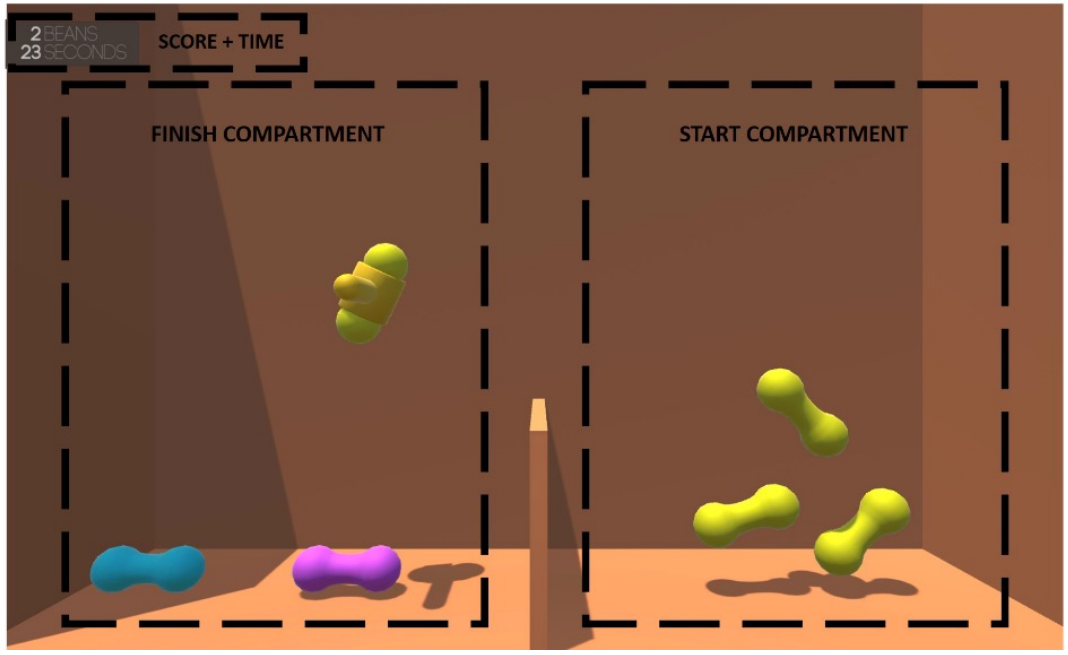
Transfer learning for sEMG

degree of freedom	original	shift	re-calibrated
pronation/supination	0.6% (0.3%)	36.2% (15.4%)	3.6% (2.5%)
flexion/extension	0.5% (0.5%)	10.9% (3.7%)	1.3% (0.9%)
open/close	0.7% (0.4%)	29.1% (6.4%)	3.8% (1.0%)

Benjamin Paaßen, Alexander Schulz, Janne Hahne, Barbara Hammer: Expectation maximization transfer learning and its application for bionic hand prostheses. Neurocomputing 298: 122-133 (2018)

UNIVERSITÄT BIELEFELD
Technische Fakultät

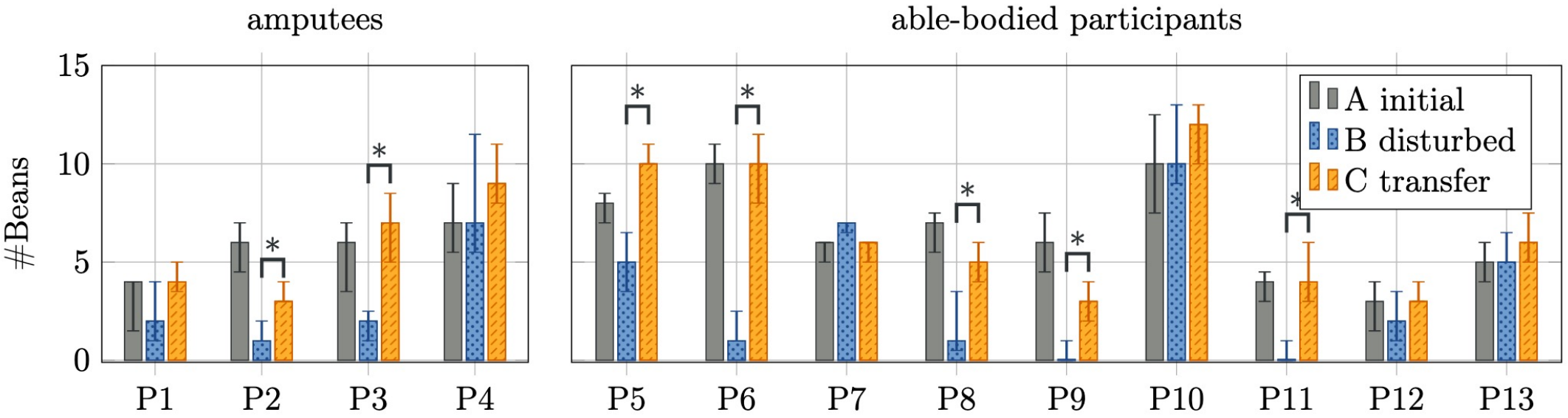
Transfer learning for sEMG



Counteracting Electrode Shifts in Upper-Limb Prosthesis Control via Transfer Learning, *Prahm C, Schulz A, Paaßen B, Schoisswohl J, Kaniusas E, Dorffner G, Hammer B, Aszmann O (2019) IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27(5): 956-962.

UNIVERSITÄT BIELEFELD
Technische Fakultät

Transfer learning for sEMG



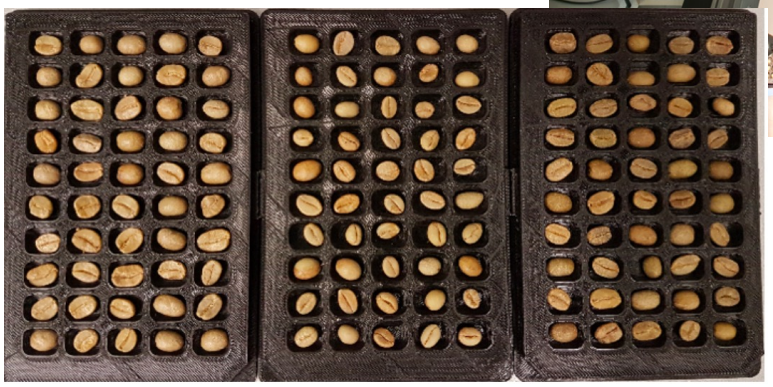
Counteracting Electrode Shifts in Upper-Limb Prosthesis Control via Transfer Learning, *Prahm C, Schulz A, Paaßen B, Schoisswohl J, Kaniusas E, Dorffner G, Hammer B, Aszmann O (2019) IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27(5): 956-962.

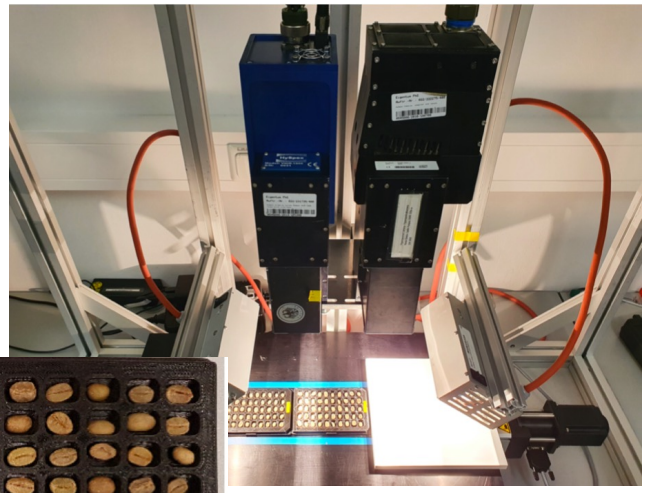
UNIVERSITÄT BIELEFELD
Technische Fakultät

Transfer learning II

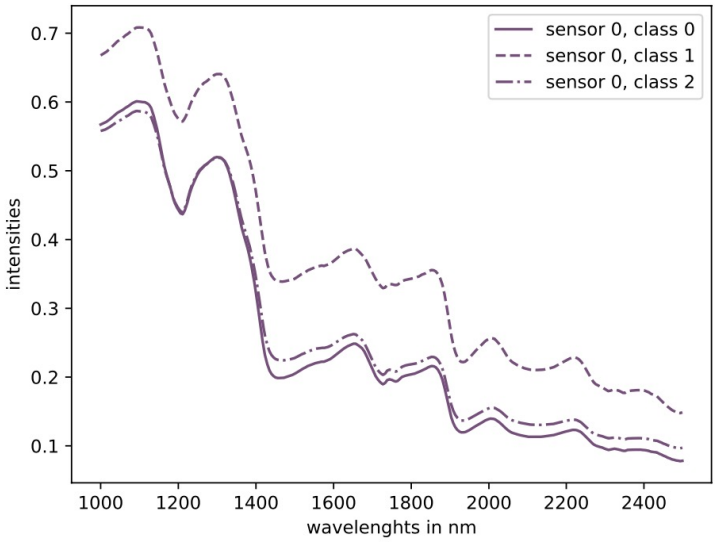
UNIVERSITÄT BIELEFELD
Technische Fakultät

Classifying coffee



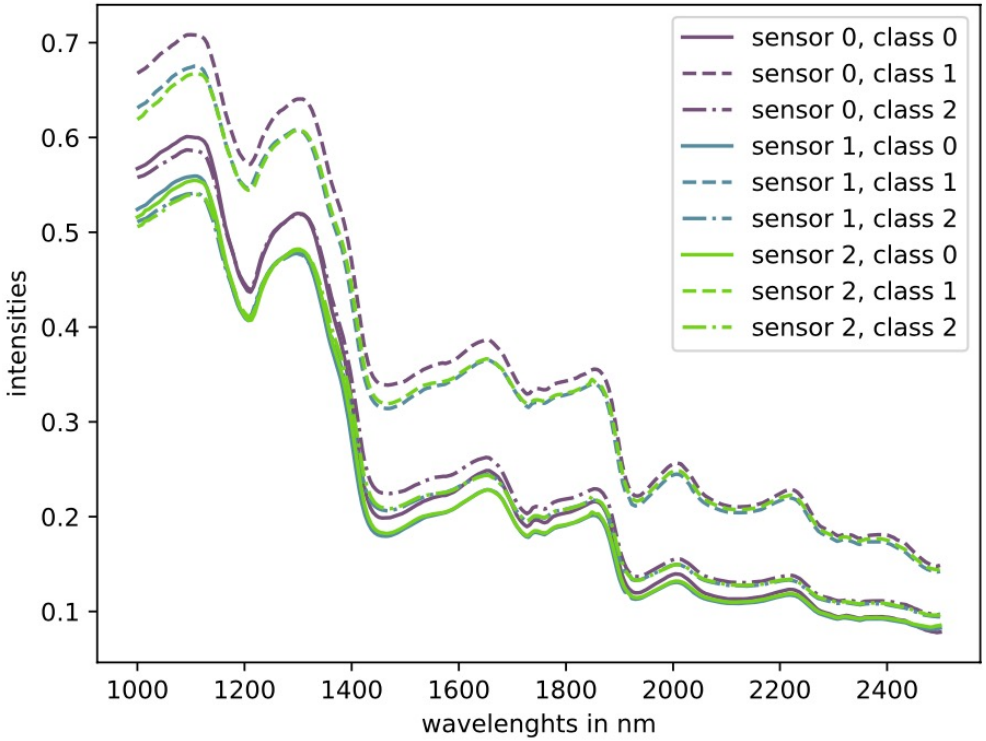


94% accuracy



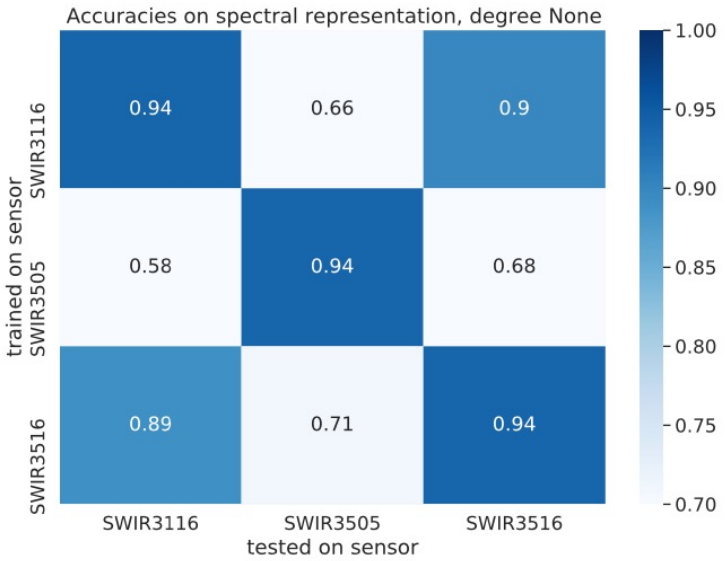
UNIVERSITÄT BIELEFELD
Technische Fakultät

Drift in hyperspectral data



UNIVERSITÄT BIELEFELD
Technische Fakultät

Drift in hyperspectral data



UNIVERSITÄT BIELEFELD
Technische Fakultät

Unsupervised transfer learning

Transfer learning:

Given a function $f: X \rightarrow Y$ which has been learned based on a probability distribution $P_{X \times Y}$

Given samples (x^i) where the distribution of the data has changed $P_{X'} \neq P_X$

Learn a classification *g* from the given data with the help of f

UNIVERSITÄT BIELEFELD
Technische Fakultät

Maximising Minimum discrepancy

Transfer learning by MMD embedding

Assumption: $P_{X'} \neq P_X$ but $P(Y'|X') = P(Y|X)$

Goal: transform X' into X

Given samples, minimize discrepancy based on universal kernel embedding

$$MMD(X, X'):$$
$$= \left| 1/n_1 \sum_{i=1..n_1} \phi(x_i) \quad - 1/\mathrm{n}_2 \sum_{i=1..n_2} \phi(x_i') \right|_H^2$$

Domain adaptation via transfer component analysisSJ Pan, IW Tsang, JT Kwok, Q Yang
IEEE Transactions on Neural Networks 22 (2), 199-210, 2011

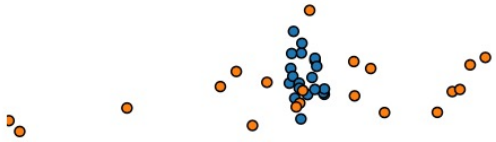
UNIVERSITÄT BIELEFELD
Technische Fakultät

Practice: Transfer learning via moment matching

	Blue		Orange
$\mathbb{E}[X]$	5.2	$\neq$	-3

UNIVERSITÄT BIELEFELD
Technische Fakultät

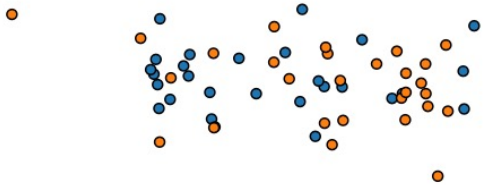
Practice: Transfer learning via moment matching



	Blue		Orange
$\mathbb{E}[X]$	0	$\approx$	$-2.6 \cdot 10^{-16}$

UNIVERSITÄT BIELEFELD
Technische Fakultät

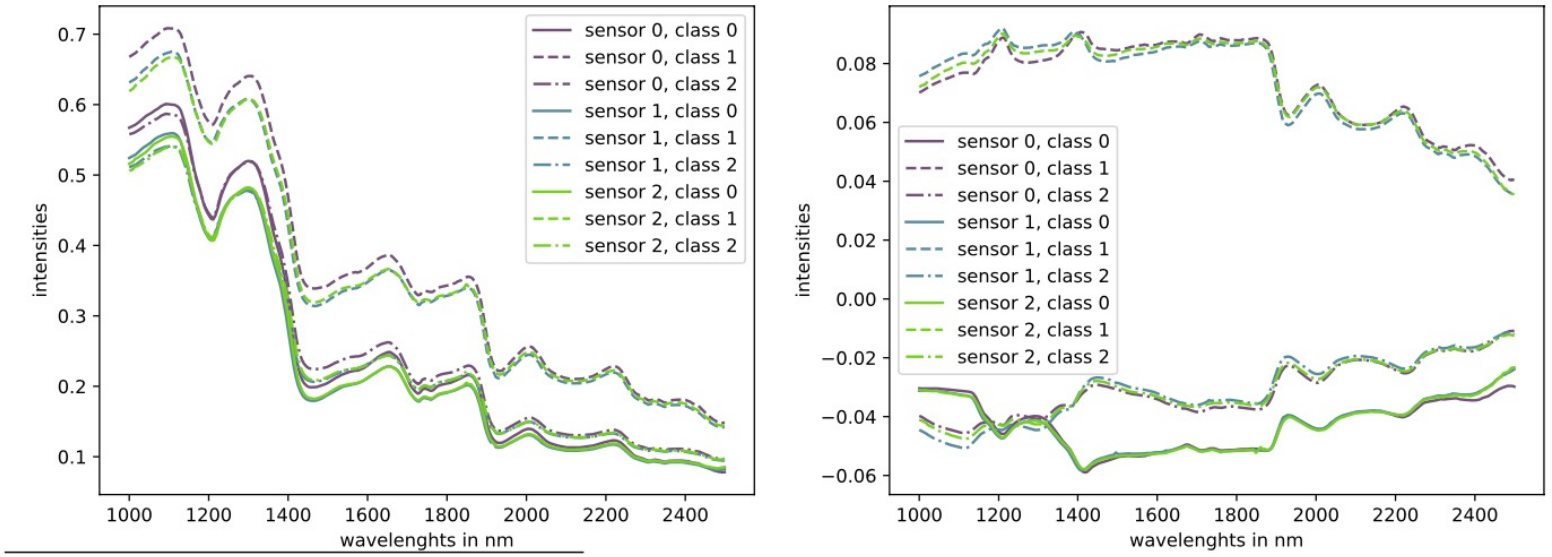
Practice: Transfer learning via moment matching



	Blue		Orange
$\mathbb{E}[X]$	$1.9 \cdot 10^{-16}$	$\approx$	$6.5 \cdot 10^{-16}$
$\mathbb{E}[X^2]$	25.0	$\approx$	24.99

UNIVERSITÄT BIELEFELD
Technische Fakultät

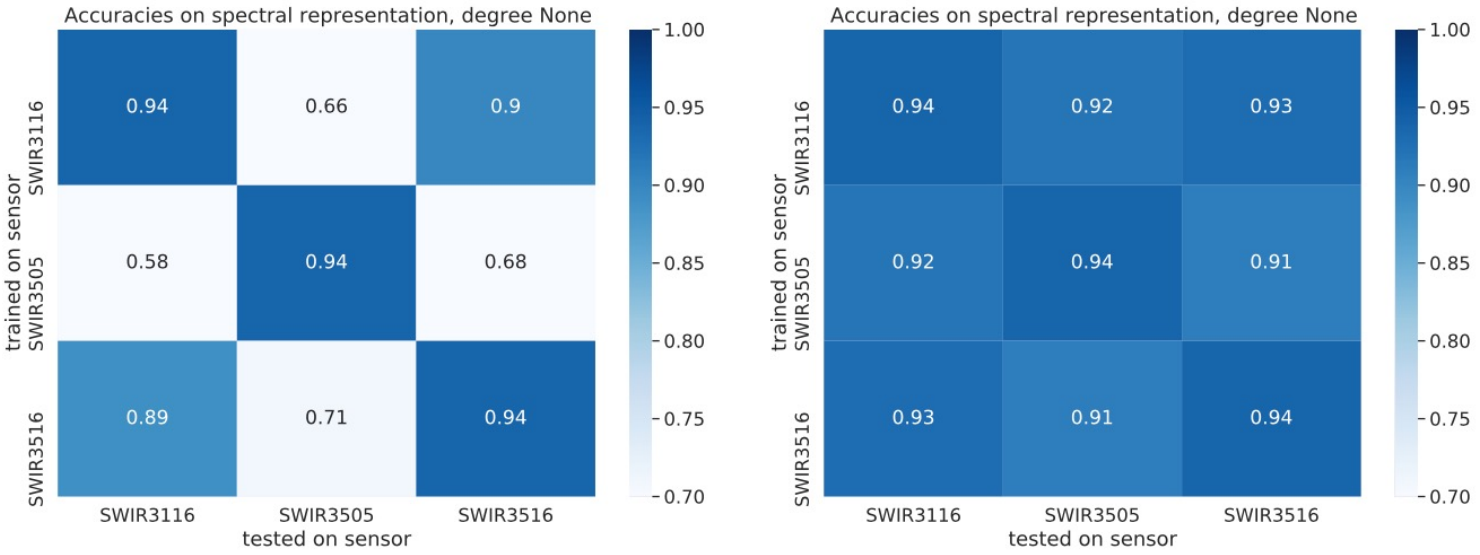
Transfer learning via moment matching for hyperspectral data



Valerie Vaquet, Patrick Menz, Udo Seiffert, Barbara Hammer, Investigating Intensity and Transversal Drift in Hyperspectral Imaging Data, ESANN2021

UNIVERSITÄT BIELEFELD
Technische Fakultät

Transfer learning via moment matching for hyperspectral data



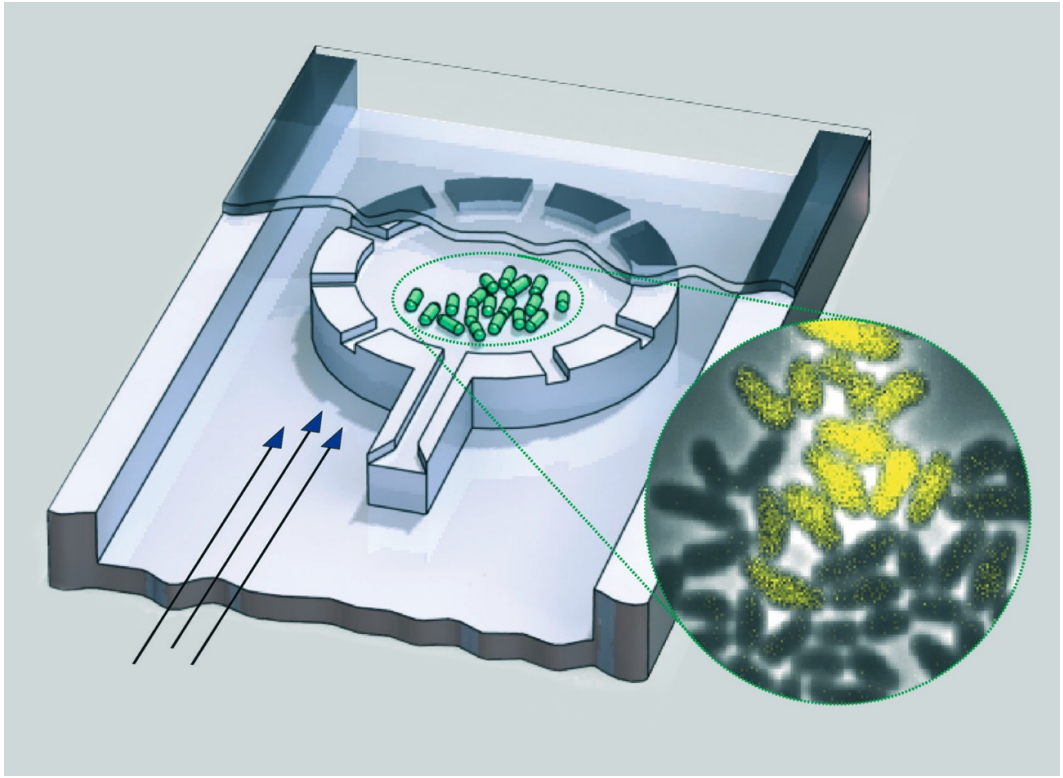
Valerie Vaquet, Patrick Menz, Udo Seiffert, Barbara Hammer, Investigating Intensity and Transversal Drift in Hyperspectral Imaging Data, ESANN2021

UNIVERSITÄT BIELEFELD
Technische Fakultät

Transfer learning III

UNIVERSITÄT BIELEFELD
Technische Fakultät

Microfluidic single-cell cultivation

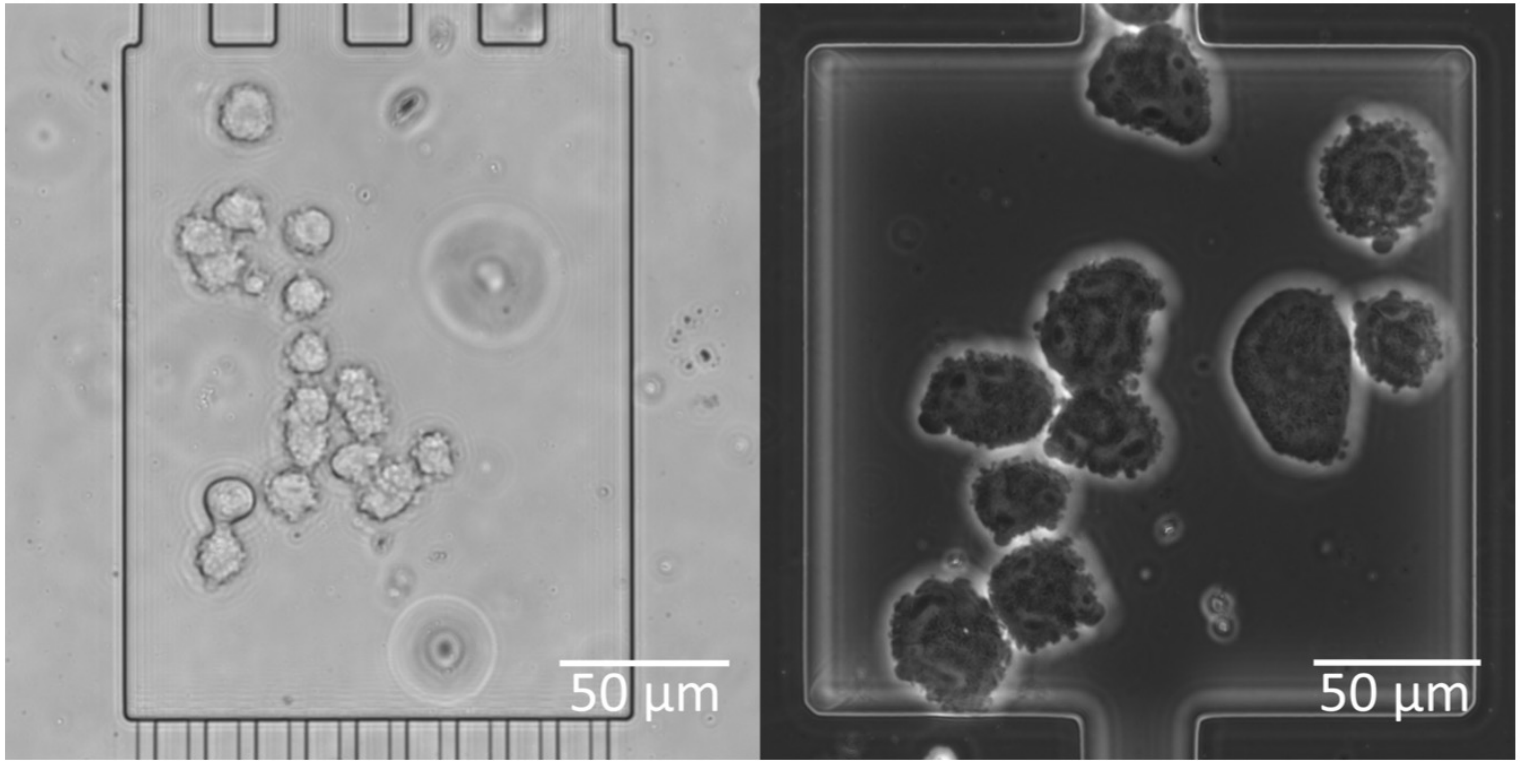


- “Lab on a chip” single-cell cultivation
- observation of single cells over time under controlled conditions
- interesting properties
  - growth rate
  - homogeneity
  - …

Key Technologies, Volume 114, Single-Cell Analysis in Microfluidic Bioreactors, Alexander Manuel Grünberger, Member of the Helmholtz Association, 2014

UNIVERSITÄT BIELEFELD
Technische Fakultät

Typical data

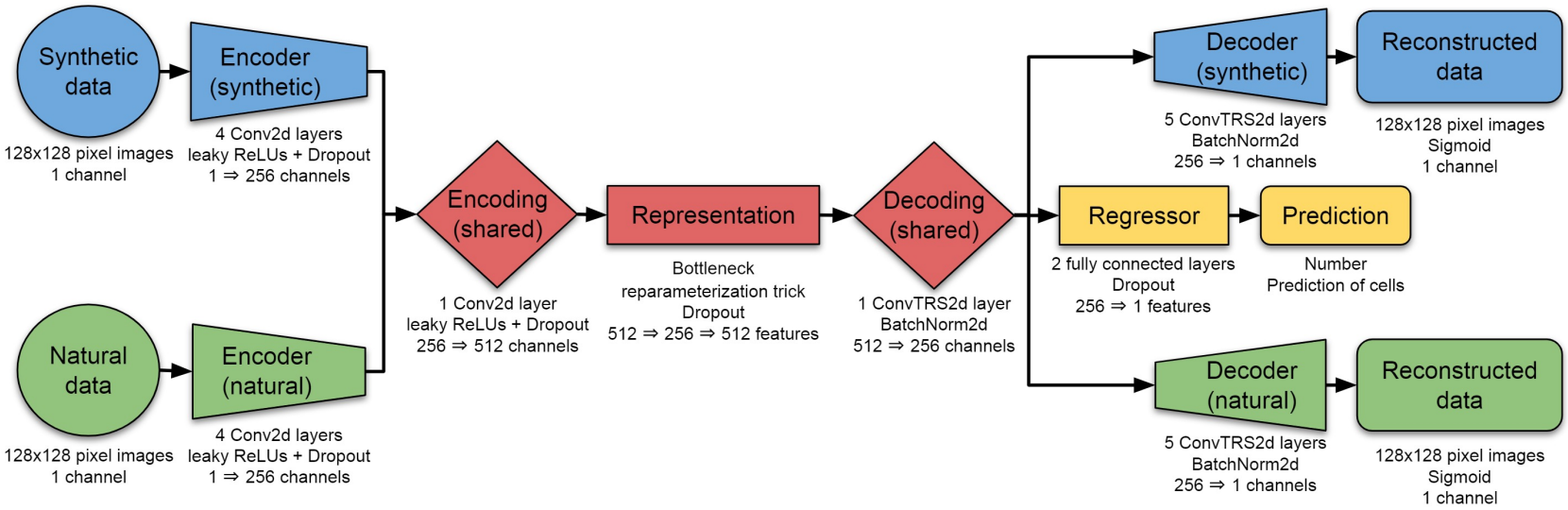


Bright field microscopy

Phase contrast microscopy

UNIVERSITÄT BIELEFELD
Technische Fakultät

Siamese autoencoder model for domain adaptation



Dominik Stallmann, Jan Philip Göpfert, Julian Schmitz, Alexander Grünberger, Barbara Hammer: Towards an automatic analysis of CHO-K1 suspension growth in microfluidic single-cell cultivation. **Bioinform**. 37(20): 3632-3639 (2021)

UNIVERSITÄT BIELEFELD
Technische Fakultät

Training objective of twin VAE

$$\mathrm{Twin}_{\mathrm{loss}}(x, l, t) =$$

$$C^t_{\mathrm{Rec}} \cdot \mathrm{Rec}(x) + C^{t,l}_{\mathrm{Reg}} \cdot \mathrm{Reg}(x, l) + C^t_{\mathcal{D}_{\mathrm{KL}}} \cdot \mathcal{D}_{\mathrm{KL}}(x)$$

reconstruction error:
squared error or cross-entropy
for reconstruction

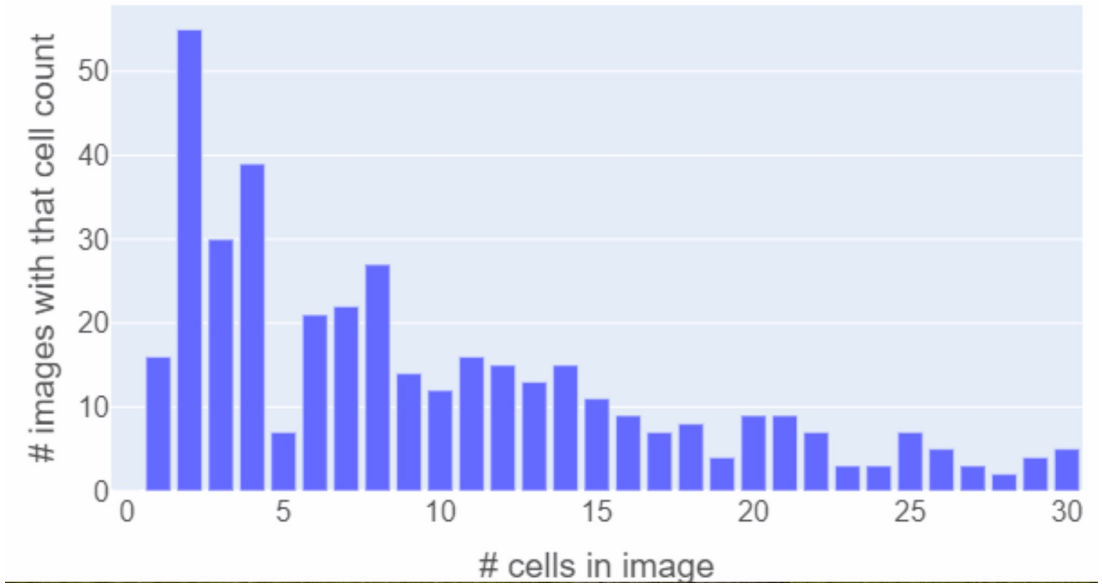
regression error
for labeling

KL error of VAE

Dominik Stallmann, Jan Philip Göpfert, Julian Schmitz, Alexander Grünberger, Barbara Hammer: Towards an automatic analysis of CHO-K1 suspension growth in microfluidic single-cell cultivation. **Bioinform**. 37(20): 3632-3639 (2021)

UNIVERSITÄT BIELEFELD
Technische Fakultät

Training data



bright field:

- ca. 5000
- 1,4% labelled

phase contrast:

- ca 11.000
- 5.9% labelled

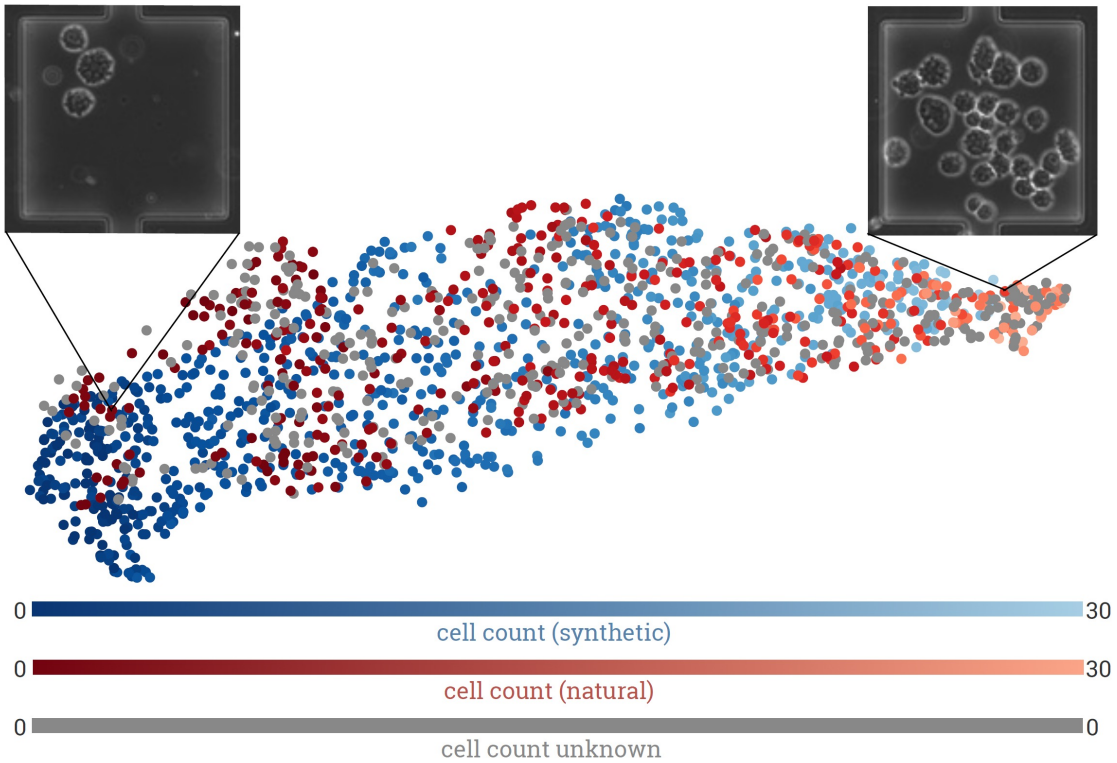
artificial:

- on the fly generation for up to 50.000 epochs

Dominik Stallmann, Jan Philip Göpfert, Julian Schmitz, Alexander Grünberger, Barbara Hammer: Towards an automatic analysis of CHO-K1 suspension growth in microfluidic single-cell cultivation. **Bioinform**. 37(20): 3632-3639 (2021)

UNIVERSITÄT BIELEFELD
Technische Fakultät

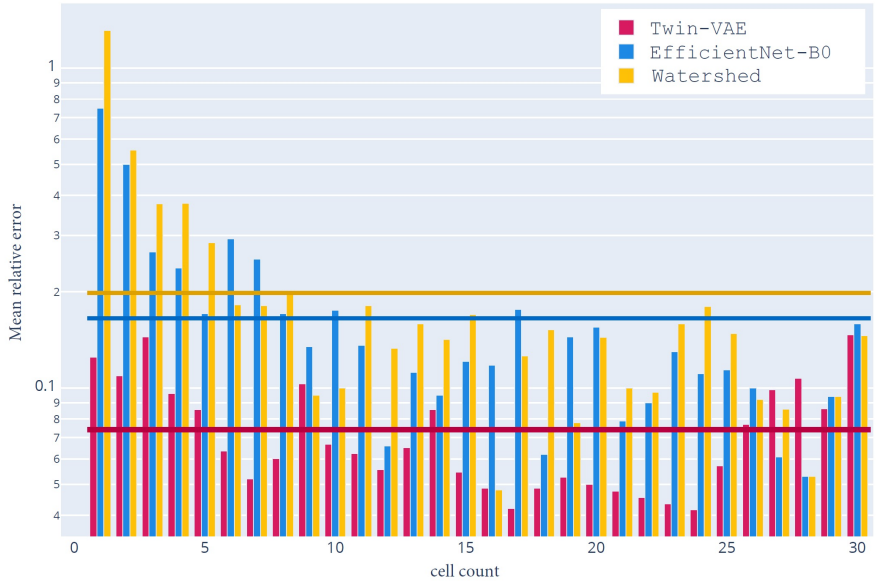
Learned representation



Dominik Stallmann, Jan Philip Göpfert, Julian Schmitz, Alexander Grünberger, Barbara Hammer: Towards an automatic analysis of CHO-K1 suspension growth in microfluidic single-cell cultivation. **Bioinform**. 37(20): 3632-3639 (2021)

UNIVERSITÄT BIELEFELD
Technische Fakultät

Results



Dominik Stallmann, Jan Philip Göpfert, Julian Schmitz, Alexander Grünberger, Barbara Hammer: Towards an automatic analysis of CHO-K1 suspension growth in microfluidic single-cell cultivation. **Bioinform**. 37(20): 3632-3639 (2021)

UNIVERSITÄT BIELEFELD
Technische Fakultät

Results including transfer learning in between domains

Method	MAE (Syn) ↓	MRE / % (Syn) ↓	Acc. / % (Syn) ↑	MAE (Nat) ↓	MRE / % (Nat) ↓	Acc. / % (Nat) ↑
			PC (phase-contrast microscopy)			
Watershed	0.94	18.00	24.00	1.66	29.00	23.10
EfficientNet-B0	4.99	79.40	5.00	1.67	25.10	23.40
EfficientNet-B1	3.53	54.50	8.70	1.78	31.90	19.30
Twin-VAE (Nat only)	n/a	n/a	n/a	1.07	20.10	39.80
Twin-VAE$_{\text{max-acc}}$	**0.09**	0.68	68.20	0.60	5.92	**57.80**
Twin-VAE$_{\text{min-dev}}$	0.14	0.73	62.10	**0.59**	**5.66**	57.00
BiT	n/a	n/a	n/a	2.20	n/a	26.13
Transfer-Twin-VAE (Nat only)	n/a	n/a	n/a	1.01	14.11	44.20
Transfer-Twin-VAE	0.15	**0.43**	**85.00**	0.81	6.48	53.70
			BF (bright-field microscopy)			
Watershed	1.92	39.00	2.00	2.39	32.00	32.00
EfficientNet-B0	6.50	67.10	4.50	1.13	17.20	33.90
EfficientNet-B1	5.25	67.50	4.40	1.21	18.50	29.00
Twin-VAE (Nat only)	n/a	n/a	n/a	0.91	13.30	23.40
Twin-VAE$_{\text{max-acc}}$	0.48	4.27	60.10	0.68	7.60	53.20
Twin-VAE$_{\text{min-dev}}$	0.52	4.63	58.20	0.63	7.31	51.90
BiT	n/a	n/a	n/a	1.03	n/a	43.10
Transfer-Twin-VAE (Nat only)	n/a	n/a	n/a	0.72	7.88	51.36
Transfer-Twin-VAE	**0.40**	**3.87**	**66.60**	**0.52**	**5.47**	**60.74**

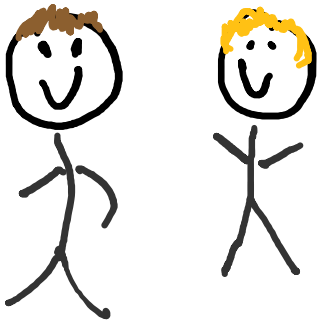
Dominik Stallmann, Jan Philip Göpfert, Julian Schmitz, Alexander Grünberger, Barbara Hammer: Towards an automatic analysis of CHO-K1 suspension growth in microfluidic single-cell cultivation. **Bioinform**. 37(20): 3632-3639 (2021)

UNIVERSITÄT BIELEFELD
Technische Fakultät

Learning with drift

UNIVERSITÄT BIELEFELD
Technische Fakultät

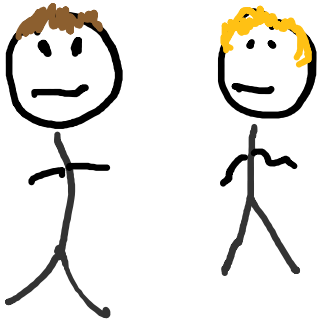
Learning with drift



UNIVERSITÄT BIELEFELD
Technische Fakultät

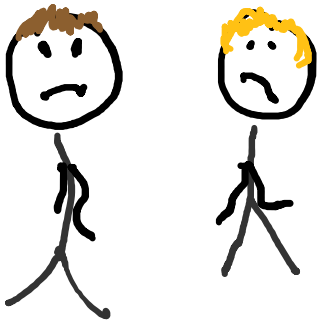
Learning with drift

2022



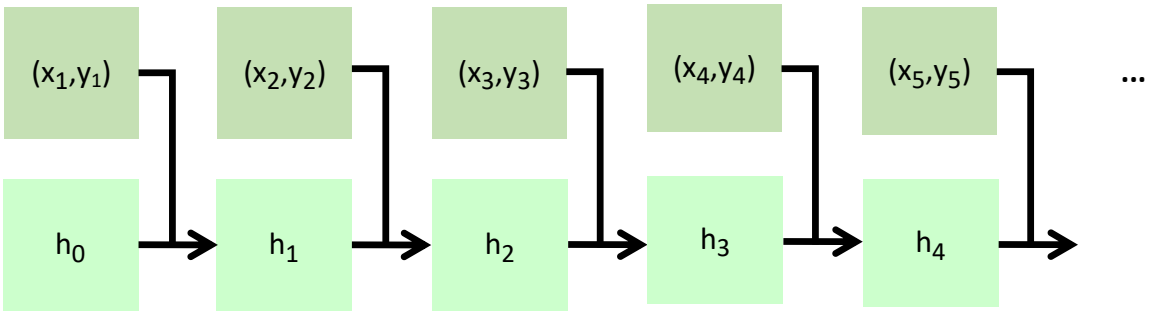
UNIVERSITÄT BIELEFELD
Technische Fakultät

Learning with drift



UNIVERSITÄT BIELEFELD
Technische Fakultät

Learning from data streams



UNIVERSITÄT BIELEFELD
Technische Fakultät

Learning from data streams

Given a **stream of training data**

$(x^1, y^1), \ldots, (x^t, y^t), \ldots \in X \times Y$

sampled w.r.t. a family of probability distributions P_t on $X \times Y$

We aim for a **learning scheme which incrementally adapts a model** $h_t : X \to Y$ based on (x^t, y^t) such that the interleaved train-test error

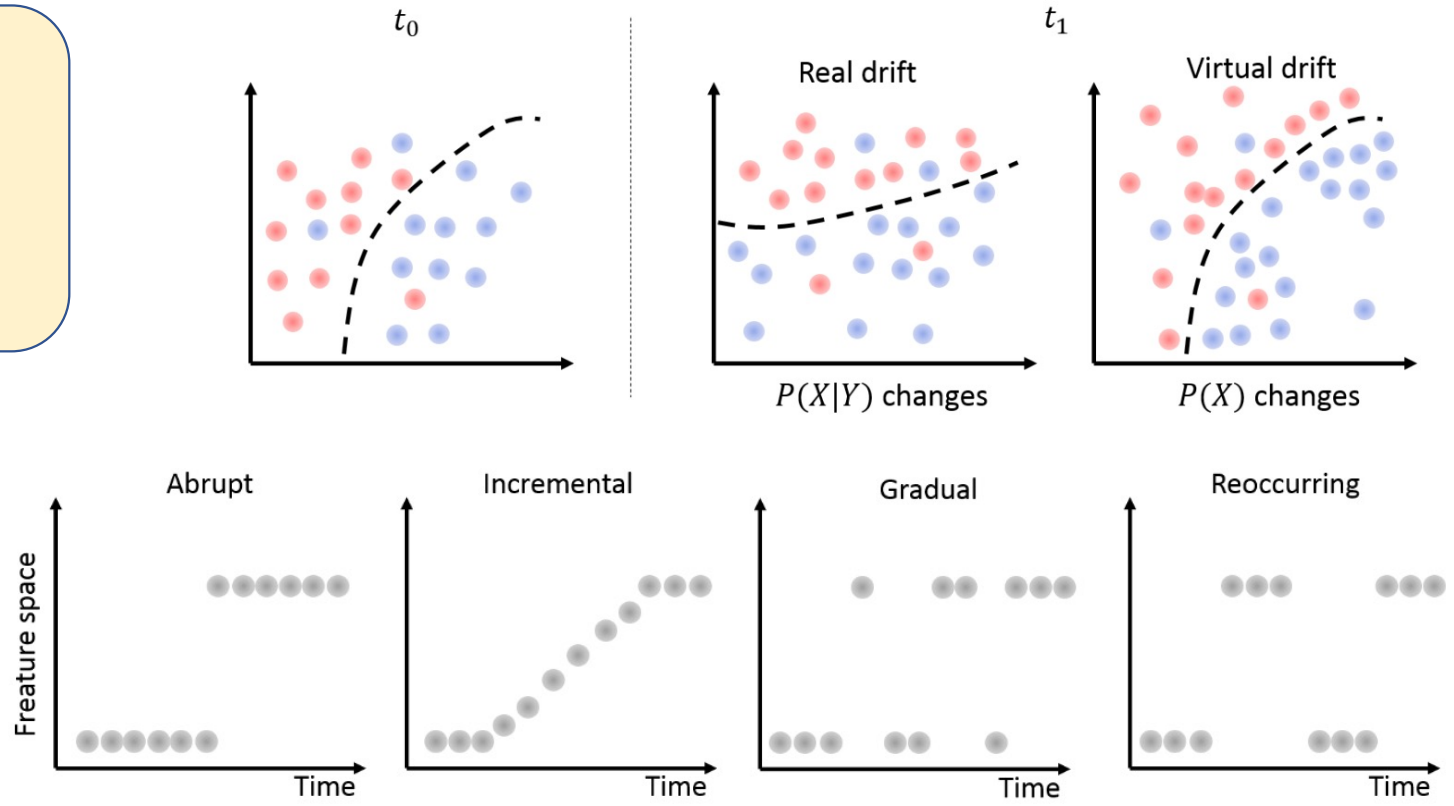
$E = \sum_t l(h_{t-1}(x_t), y_t)$ is minimized.

UNIVERSITÄT BIELEFELD
Technische Fakultät

Drift

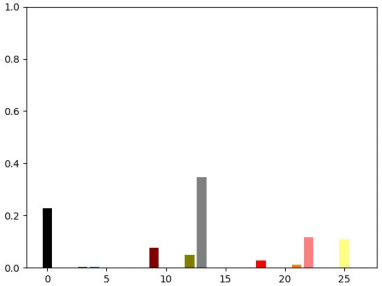
Drift is present if there exist time points $t_1 \neq t_2$ such that

$$P_{t_1} \neq P_{t_2}$$



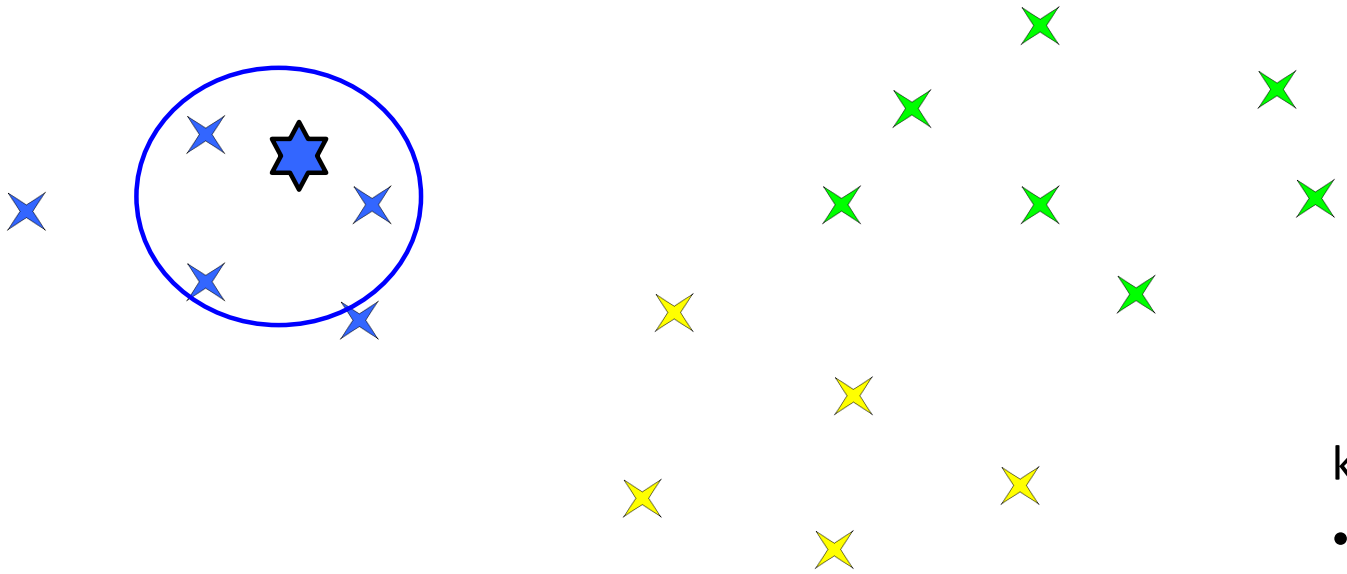
UNIVERSITÄT BIELEFELD
Technische Fakultät

Drift



UNIVERSITÄT BIELEFELD
Technische Fakultät

k-NN: basic incremental model

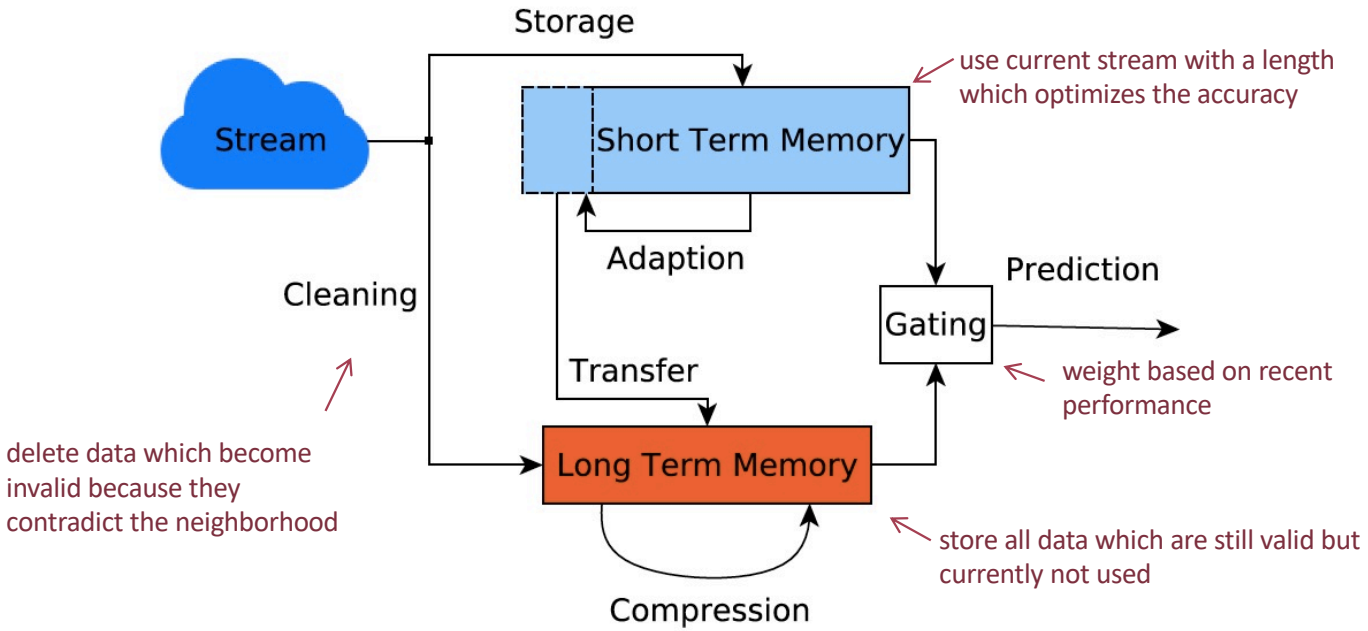


k-NN for drifting data:

- how much data to store?

UNIVERSITÄT BIELEFELD
Technische Fakultät

Self-adjusting memory (SAM-kNN)



Parameters:

- size of STM
- data points in LTM
- weights of gating

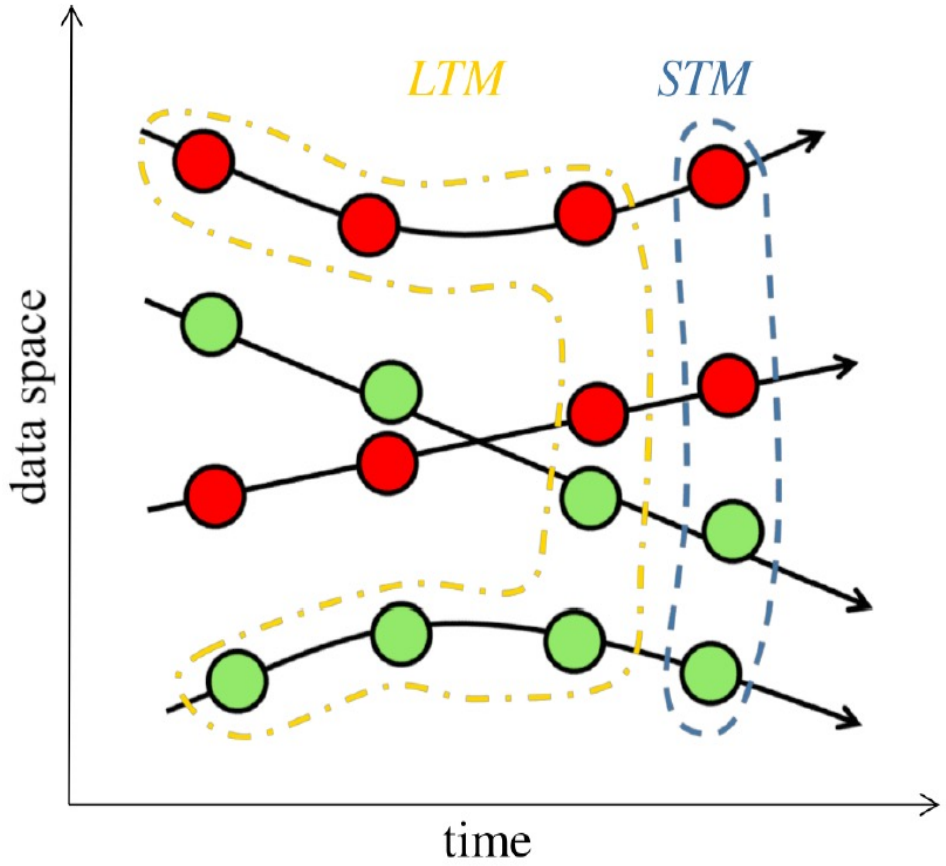
Meta-parameters:

- min size of STM
- max size of STM and LTM
- k of k-NN

Viktor Losing, Barbara Hammer, Heiko Wersing: Tackling heterogeneous concept drift with the Self-Adjusting Memory (SAM). Knowl. Inf. Syst. 54(1): 171-201 (2018), code: https://github.com/vlosing/SAMkNN or within RIVER: https://riverml.xyz/latest/ as SAMKNNClassifier

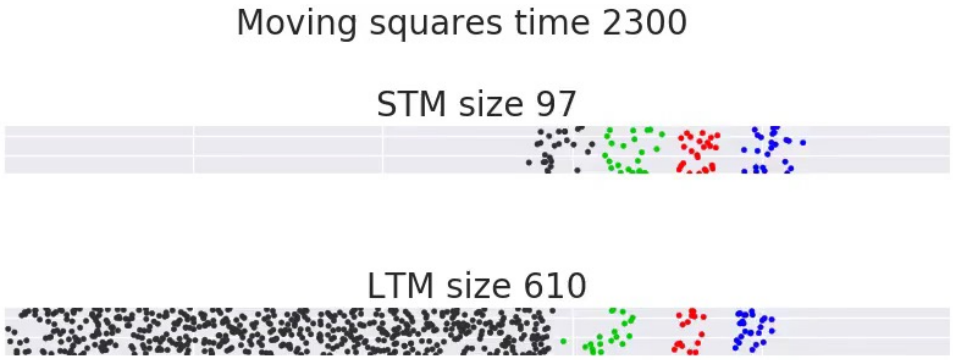
UNIVERSITÄT BIELEFELD
Technische Fakultät

Role of memory



UNIVERSITÄT BIELEFELD
Technische Fakultät

SAM-kNN – example memory



Viktor Losing, Barbara Hammer, Heiko Wersing: Tackling heterogeneous concept drift with the Self-Adjusting Memory (SAM). Knowl. Inf. Syst. 54(1): 171-201 (2018), code: https://github.com/vlosing/SAMkNN or within RIVER: https://riverml.xyz/latest/ as SAMKNNClassifier

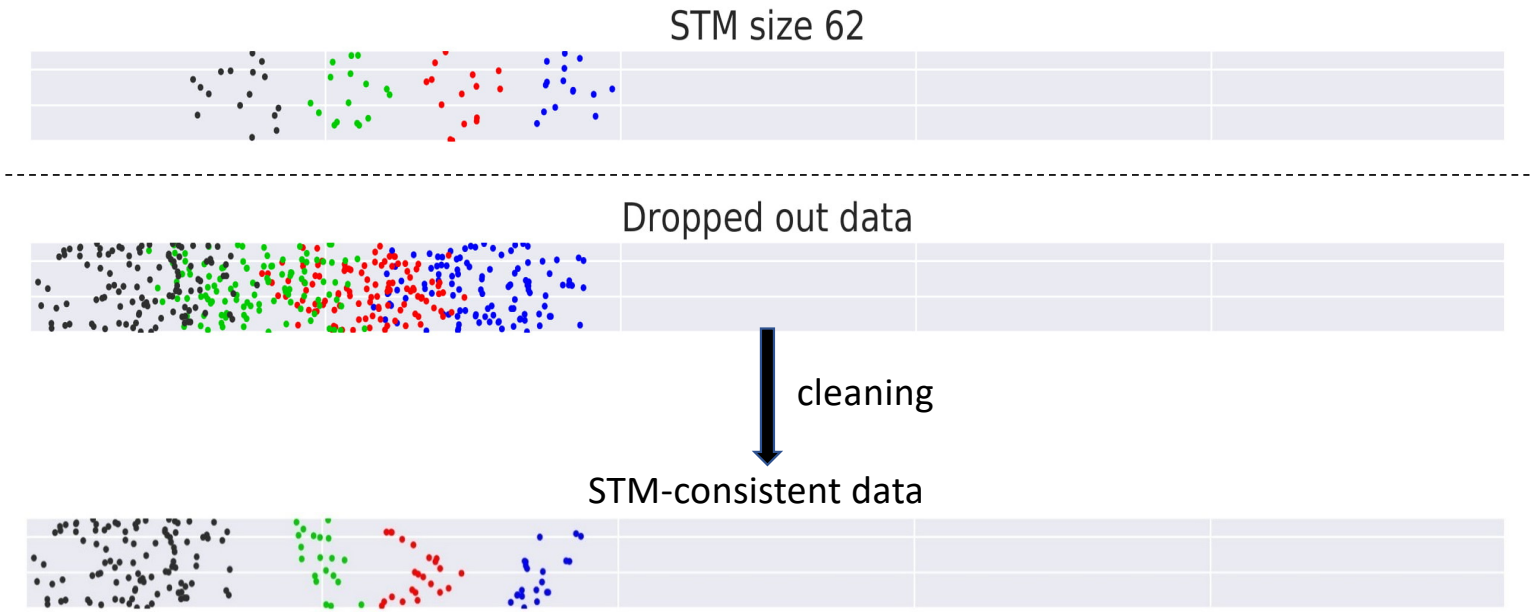
UNIVERSITÄT BIELEFELD
Technische Fakultät

STM adaptation

UNIVERSITÄT BIELEFELD
Technische Fakultät

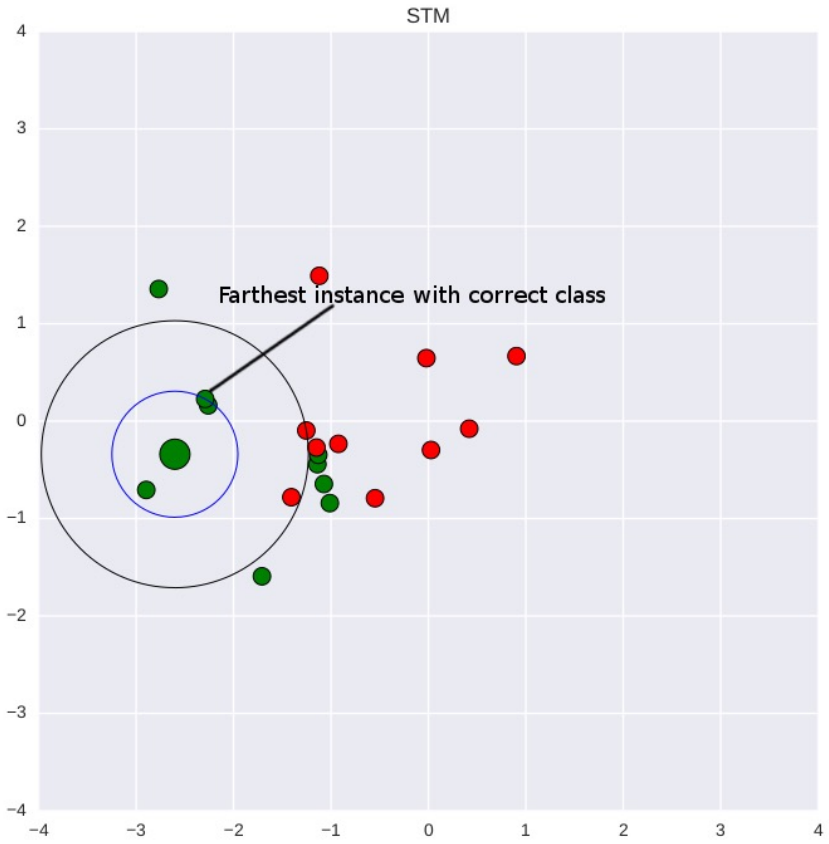
LTM

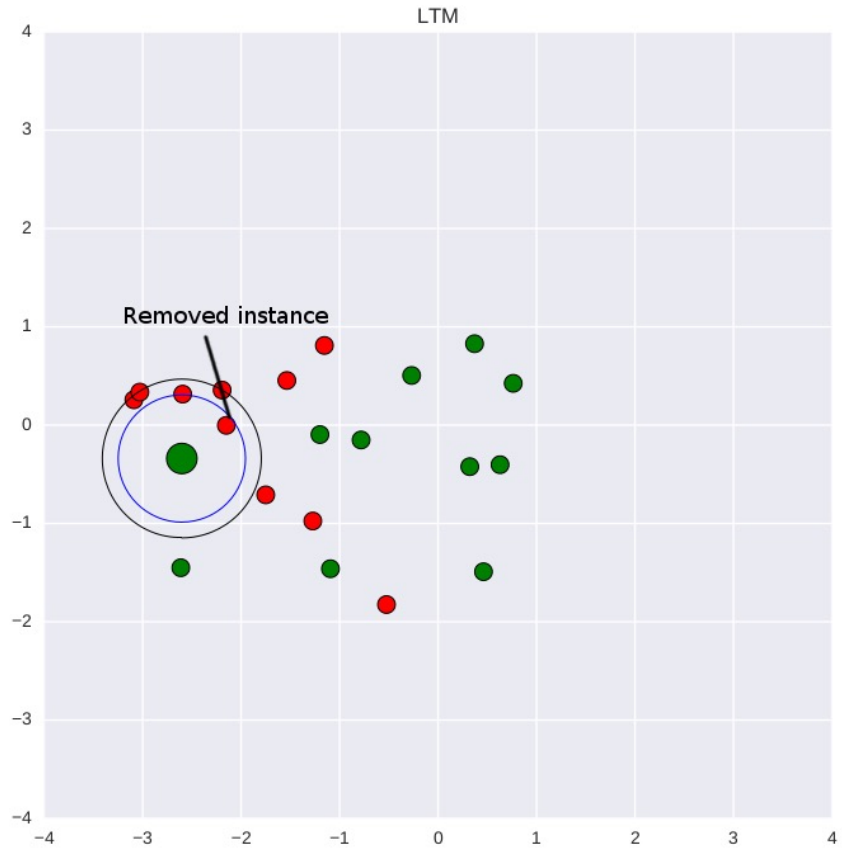
Transfer consistent data to LTM



UNIVERSITÄT BIELEFELD
Technische Fakultät

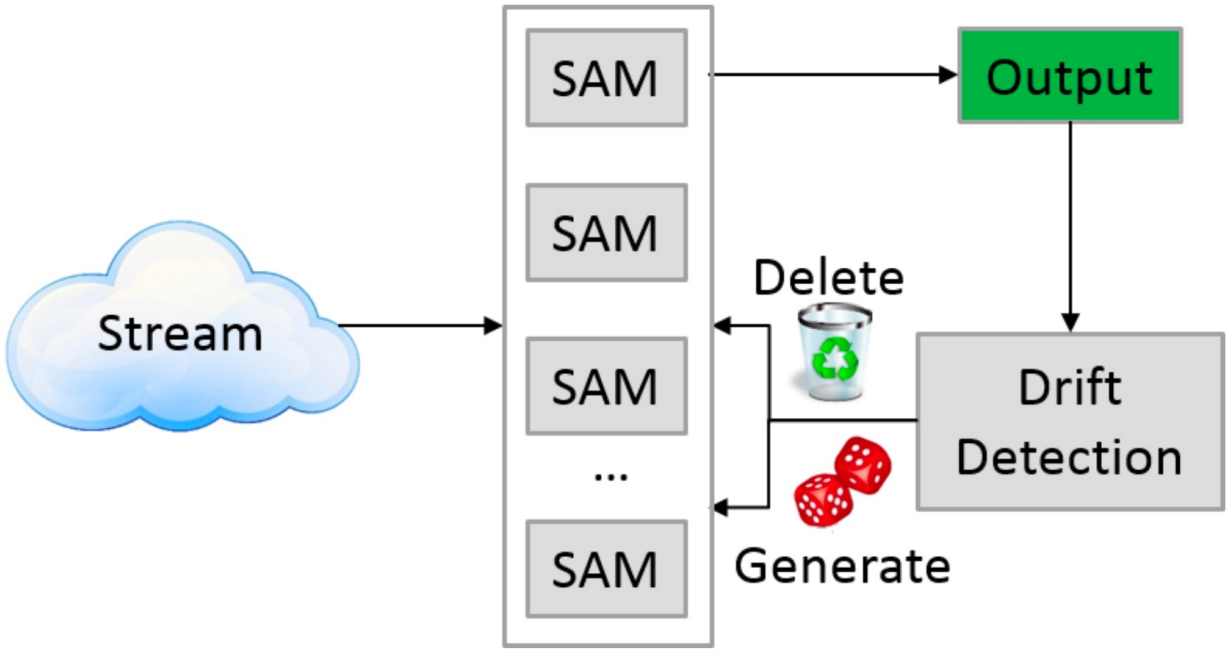
LTM





UNIVERSITÄT BIELEFELD
Technische Fakultät

Self-adjusting memory ensemble (SAME)



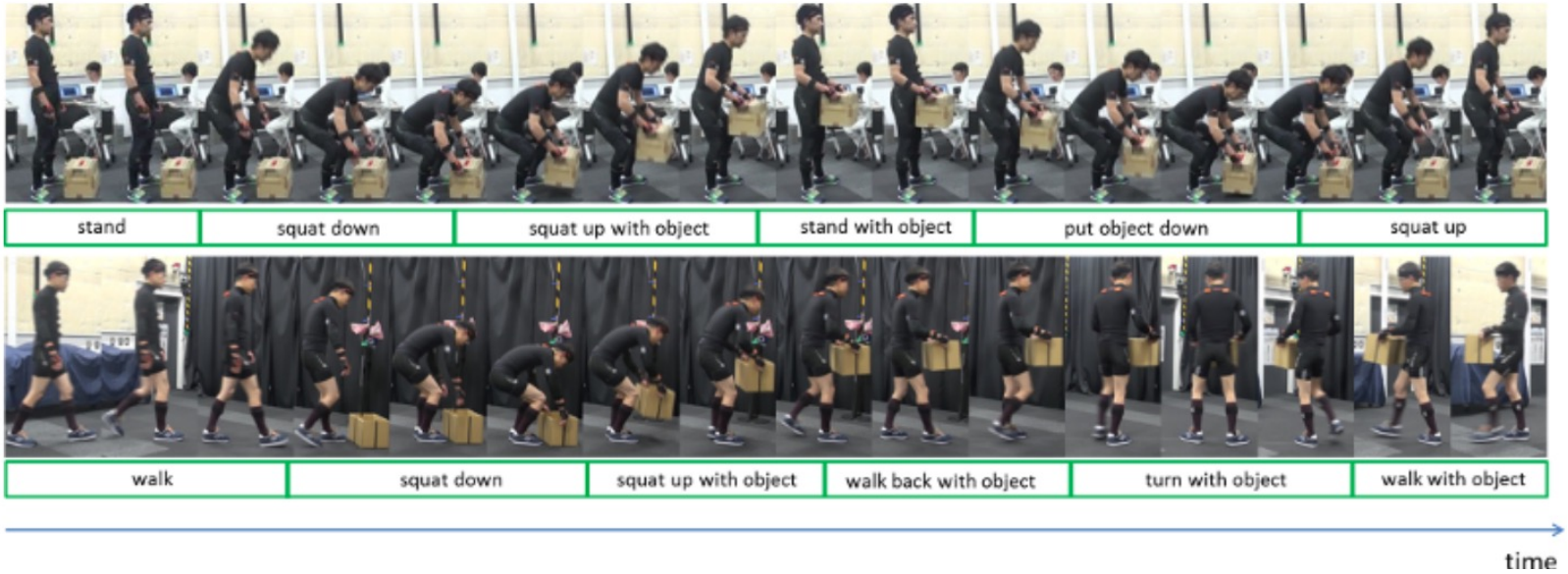
Viktor Losing, Barbara Hammer, Heiko Wersing, Albert Bifet:
Randomizing the Self-Adjusting Memory for Enhanced Handling of Concept Drift. IJCNN 2020: 1-8

Data set	VFDT	SAM	ARF	LVGB	SAM-E
SEA Concepts	15.16	13.22	11.68±0.06	**11.68**±0.07	12.28±0.07
Rot. Hyperplane	15.02	15.22	17.35±0.15	12.73±0.02	**12.49**±0.71
Moving RBF	66.27	12.10	34.02±0.17	45.62±0.15	**11.86**±0.09
Inter. RBF	74.71	3.27	**2.68**±0.04	10.08±0.94	3.30±0.01
Moving Squares	66.73	2.64	36.84±1.49	11.74±0.03	**2.47**±0.25
Transient Chessb.	45.24	11.26	26.30±0.17	14.69±6.22	**10.30**±0.09
Random Tree	10.36	37.05	8.71±1.49	**3.93**±0.09	32.72±0.77
LED-Drift	26.30	45.99	27.39±0.33	**26.13**±0.02	35.48±2.61
Mixed Drift	55.42	12.27	19.87±0.06	25.97±0.10	**11.58**±0.02
Poker	25.88	16.86	19.23±0.17	17.93±0.40	**8.79**±0.44
Artificial ∅	40.11	16.99	20.41	18.05	**14.13**
Outdoor	42.68	11.58	29.70±2.03	39.28±0.25	**9.25**±0.29
Weather	26.49	22.31	21.87±0.46	22.18±0.08	**21.41**±0.16
Electricity	29.00	17.58	21.13±0.50	17.58±0.18	**16.36**±0.19
Rialto	76.19	18.27	24.08±0.10	40.46±0.07	**15.80**±0.16
Airline	34.94	39.84	**34.20**±0.11	36.89±0.02	35.51±0.16
Cover Type	21.85	5.76	8.33±0.03	8.54±0.06	**4.69**±0.36
PAMAP	1.22	0.02	0.03±0.00	0.11±0.01	**0.02**±0.00
SPAM	19.09	7.00	8.18±0.42	7.35±0.31	**5.61**±0.23
KDD99	0.10	0.01	0.03±0.00	0.03±0.00	**0.01**±0.00
Real world ∅	27.95	13.60	16.39	19.16	**12.07**
Overall ∅	34.35	15.38	18.51	18.57	**13.15**
Overall ∅ rank	4.47	2.76	3.00	3.08	**1.68**

Nemenyi significance: SAM-E $\succ$ VFDT

UNIVERSITÄT BIELEFELD
Technische Fakultät

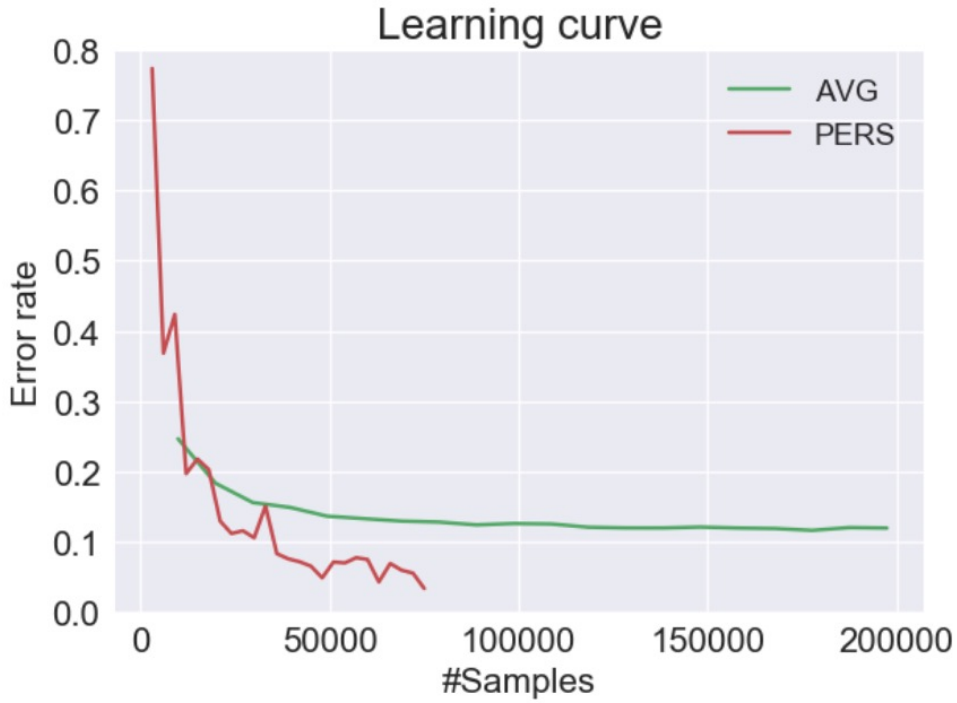
Personalized prognosis of motions



Viktor Losing, Taizo Yoshikawa, Martina Hasenjäger, Barbara Hammer, Heiko Wersing:
Personalized Online Learning of Whole-Body Motion Classes using Multiple Inertial Measurement Units. ICRA 2019: 9530-9536

UNIVERSITÄT BIELEFELD
Technische Fakultät

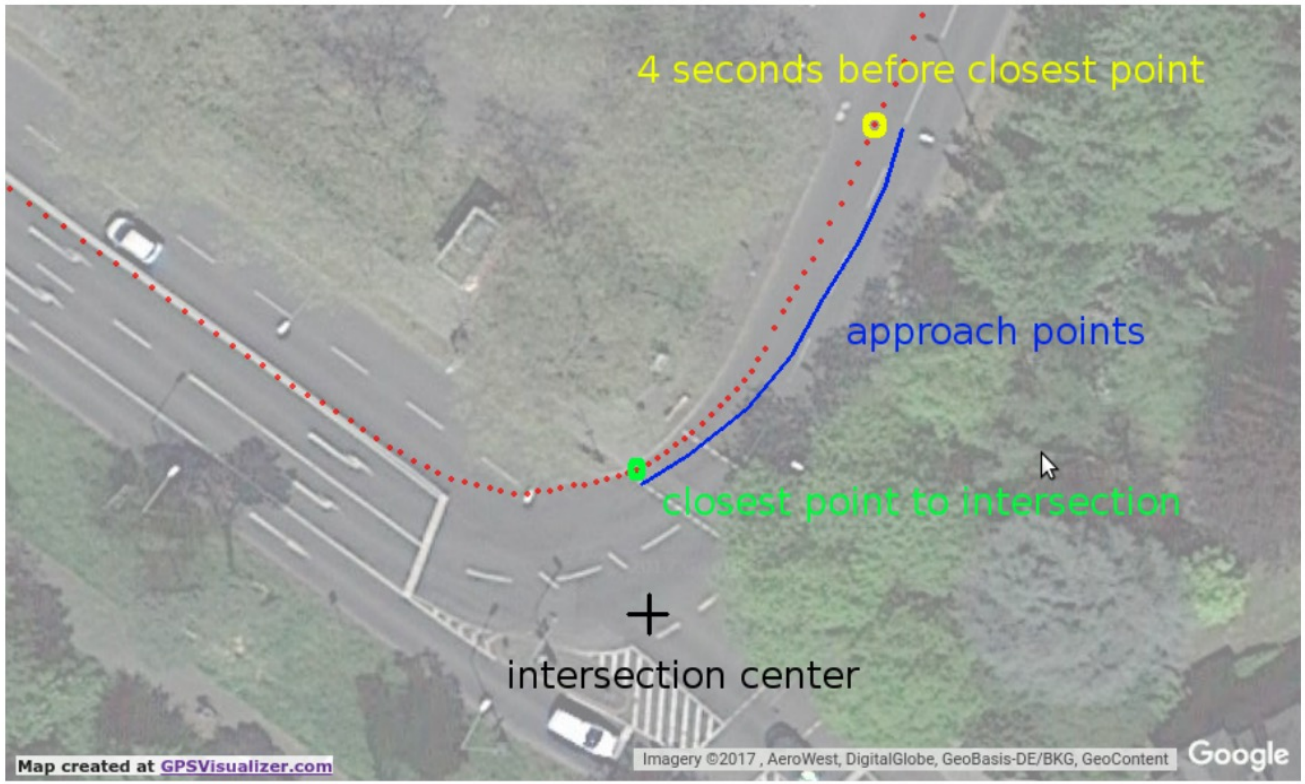
Personalized prognosis of motions



Viktor Losing, Taizo Yoshikawa, Martina Hasenjäger, Barbara Hammer, Heiko Wersing:
Personalized Online Learning of Whole-Body Motion Classes using Multiple Inertial Measurement Units. ICRA 2019: 9530-9536

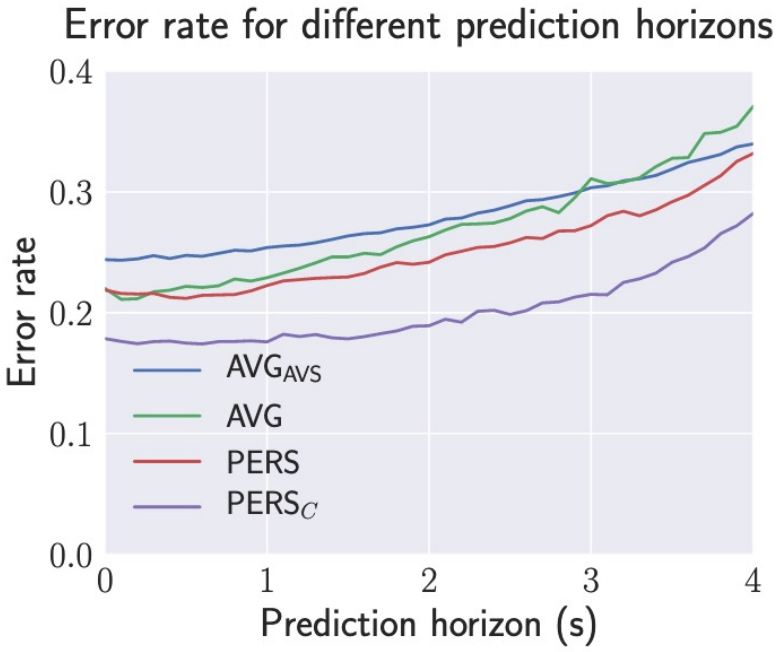
UNIVERSITÄT BIELEFELD
Technische Fakultät

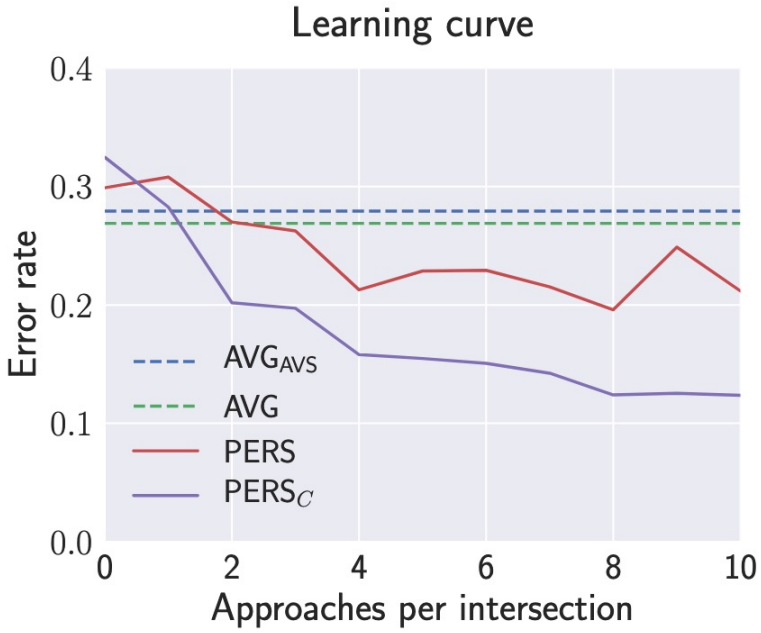
Personalized assistant for crossings



Losing, Hammer, Wersing
"Personalized Maneuver Prediction at Intersections", ITSC 2017

Personalized assistant for crossings





Losing, Hammer, Wersing
"Personalized Maneuver Prediction at Intersections", ITSC 2017

UNIVERSITÄT BIELEFELD
Technische Fakultät

More on drift

UNIVERSITÄT BIELEFELD
Technische Fakultät

What is drift?

Drift: data are drawn from a probability distribution P_t which is **not constant** with t

… but we cannot observe P_t

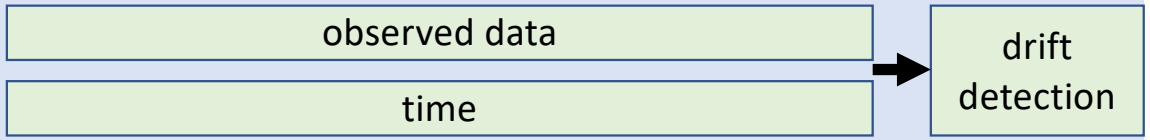
UNIVERSITÄT BIELEFELD
Technische Fakultät

Notions of drift

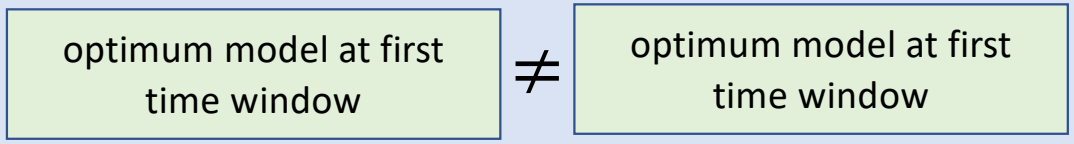
Drift: data are drawn from a probability distribution P_t which is **not constant** with t

Drift as change of (unobservable) distribution

Dependency of observations and time:



Machine Learner's drift:



Towards non-parametric drift detection via Dynamic Adapting Window Independence Drift Detection (DAWIDD), Fabian Hinder, André Artelt, Barbara Hammer, ICML2020

Definition

A *drift process* (p_t, P_T) is a probability measure P_T on $[0,1]$ together with a collection of probability measures p_t on $\mathbb{R}^d$ for all $t \in [0,1]$, such that $t \mapsto p_t(A)$ is measurable for every measurable $A \subset \mathbb{R}^d$, i.e. p_t is a Markov kernel.

Let (p_t, P_T) be a drift process. We say that p_t has *drift* iff $p_t = p_s$ does not hold for P_T-almost all $t, s \in [0,1]$.

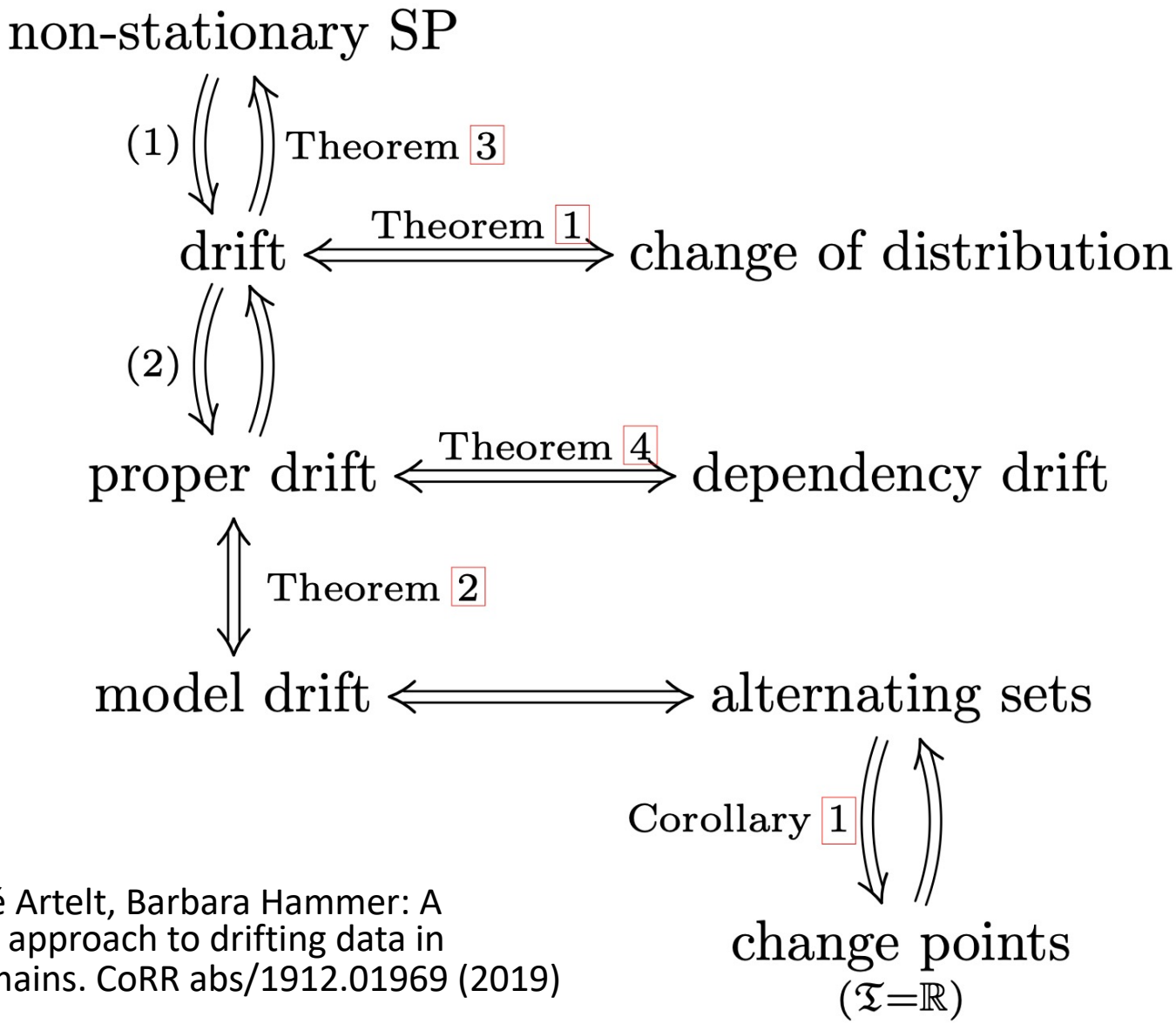
Definition

Let (p_t, P_T) be a drift process and let $(X, T) \sim p_t \otimes P_T$ a pair of random variables. We say that p_t has *dependency drift* iff X and T are statistically dependent, i.e. are not independent random variables.

Definition

We say that a drift process (p_t, P_T) has *model drift* iff there exists measurable sets $A, B \subset [0,1]$ with $P_T(A), P_T(B) > 0$, such that $p_A \neq p_B$, with $p_A = P_T(A)^{-1} \int_A p_t(\cdot) P_T(\mathrm{d}t)$ and analogous for p_B.

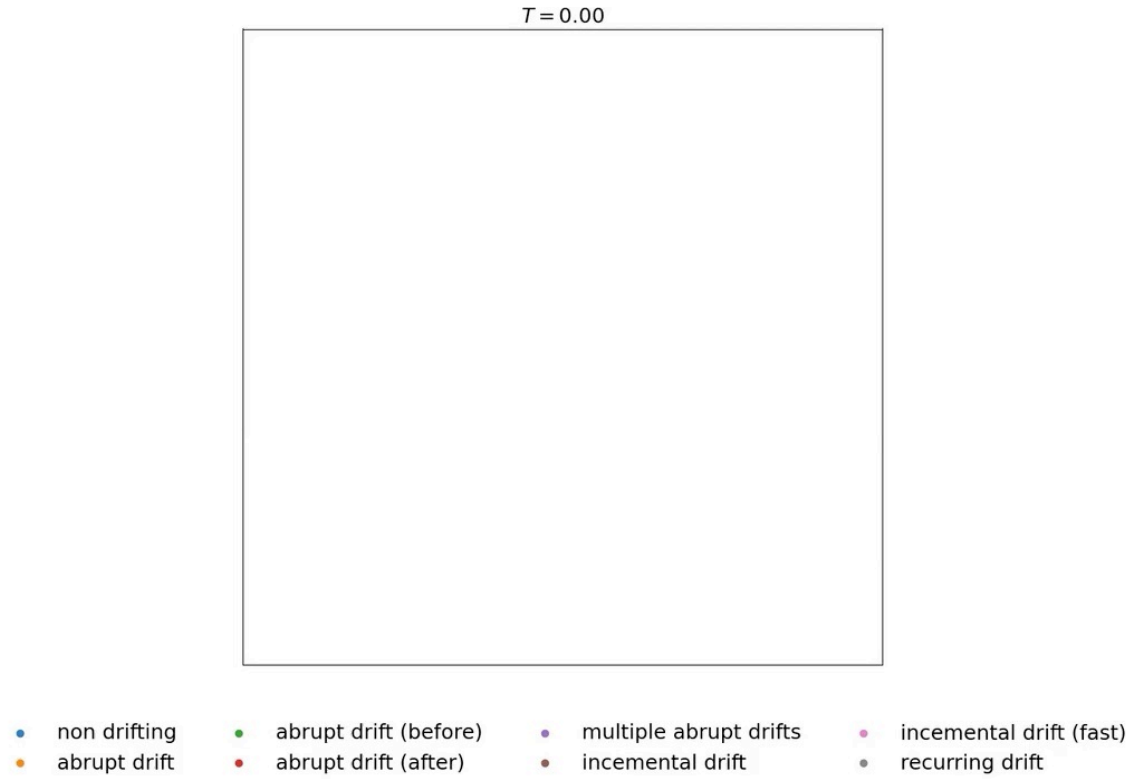
UNIVERSITÄT BIELEFELD
Technische Fakultät



Fabian Hinder, André Artelt, Barbara Hammer: A probability theoretic approach to drifting data in continuous time domains. CoRR abs/1912.01969 (2019)

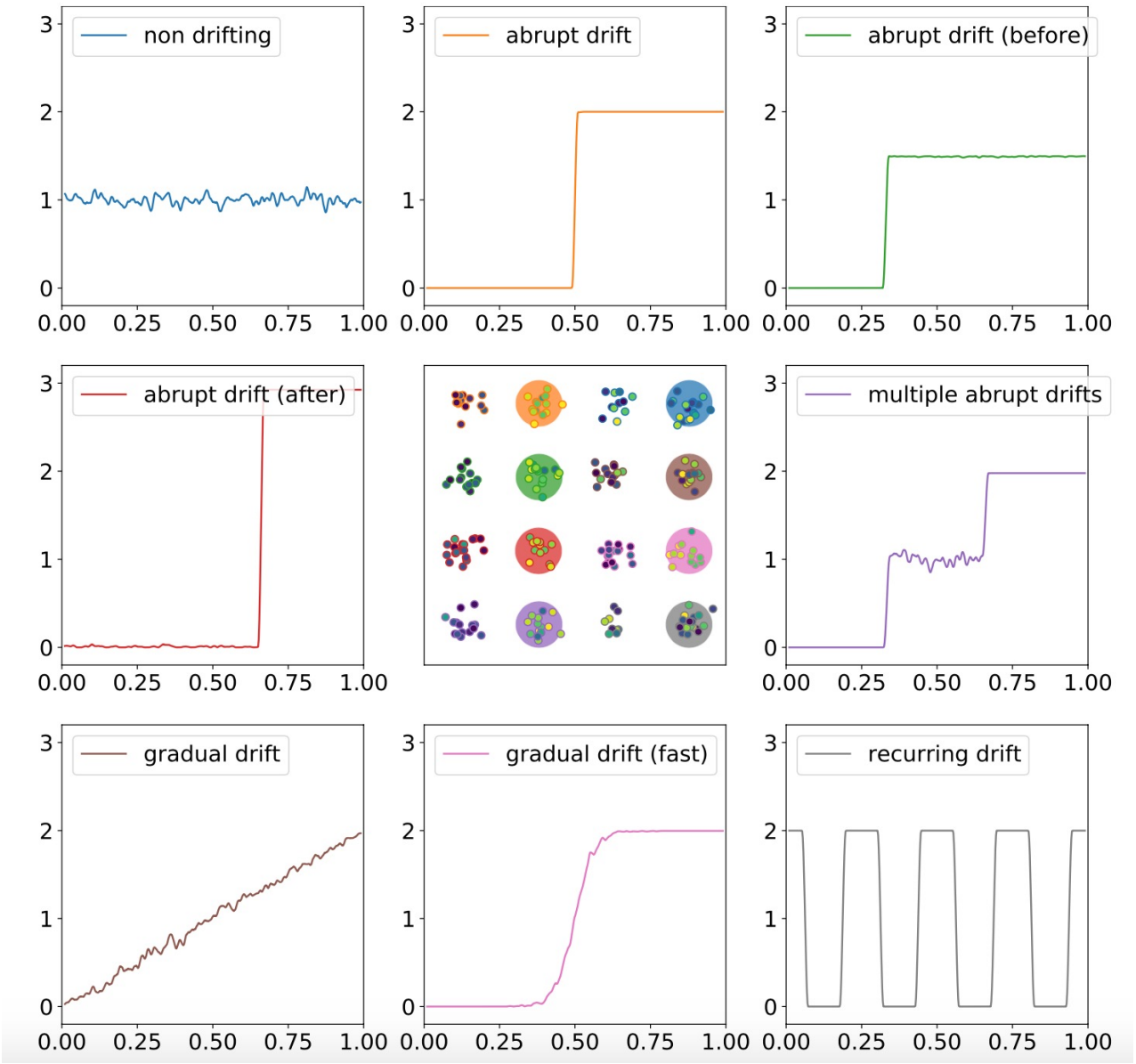
UNIVERSITÄT BIELEFELD
Technische Fakultät

Where is drift?



UNIVERSITÄT BIELEFELD
Technische Fakultät

Where is drift?



UNIVERSITÄT BIELEFELD
Technische Fakultät

Drift segmentation

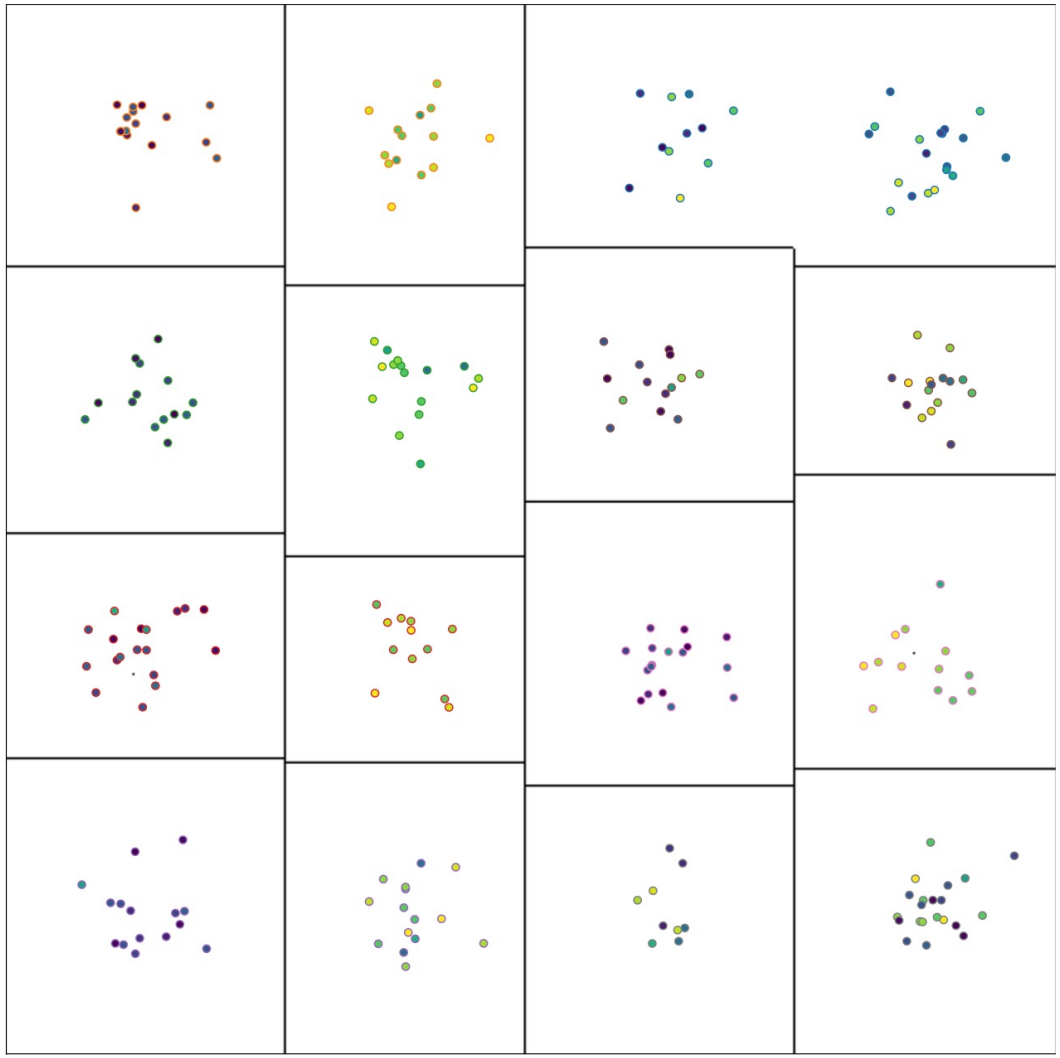
Find segmentation function $L: X \rightarrow \mathbb{N}$ such that

$$L(x) = L(x') \Rightarrow P(T|X = x) = P(T|X = x')$$

Algorithm:
iteratively split data along the axis into subsets L_i such that the homogeneity within one class is minimized / heterogeneity between classes is maximized

UNIVERSITÄT BIELEFELD
Technische Fakultät

Drift segmentation



UNIVERSITÄT BIELEFELD
Technische Fakultät

Drift localization

cpt	n	Kolmogorov	k-NN	LDD-DSI	kdq-Tree
9	0	$\mathbf{0.87(\pm 0.09)}$	$\mathbf{0.86(\pm 0.07)}$	$0.60(\pm 0.03)$	$0.78(\pm 0.11)$
9	1	$\mathbf{0.86(\pm 0.11)}$	$0.75(\pm 0.07)$	$0.49(\pm 0.06)$	$0.70(\pm 0.09)$
18	0	$0.73(\pm 0.09)$	$\mathbf{0.78(\pm 0.05)}$	$0.60(\pm 0.03)$	$0.72(\pm 0.08)$
18	1	$\mathbf{0.74(\pm 0.09)}$	$0.69(\pm 0.04)$	$0.48(\pm 0.06)$	$0.66(\pm 0.06)$
18	5	$\mathbf{0.71(\pm 0.10)}$	$0.58(\pm 0.01)$	$0.37(\pm 0.02)$	$0.48(\pm 0.05)$

F.Hinder, B.Hammer, Concept drift segmentation via Kolmogorov trees, ESANN21

UNIVERSITÄT BIELEFELD
Technische Fakultät

Conclusions

- GMLVQ → https://github.com/si-cim/prototorch, https://github.com/si-cim/protoflow
- transfer learning, not only for deep networks
- learning with drift → https://riverml.xyz/latest/ : SAMkNNclassifier
- drift segmentation and many open challenges

UNIVERSITÄT BIELEFELD
Technische Fakultät

Conclusion

joint with:

Alexander Schulz, Fabian Hinder, Johannes Brinkrolf, Michael Biehl, Petra Schneider, Lydia Fischer, Heiko Wersing, Frank-Michael Schleif, Javier Gonzalez Monroy, Javier González Jiménez, José-Luis Blanco-Claraco, Nicolai Petkov, Viktor Losing, Heiko Wersing, Jan Göpfert, Valerie Vaquet, Fabian Hinder, Andre Artelt, Taozo Yoshikawa, Martina Hasenjäger, Albert Bifet

VolkswagenStiftung
Bundesministerium für Bildung und Forschung
DFG
HRI Honda Research Institute EU
Bundesministerium für Wirtschaft und Energie
Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen