Dataset shift and its impact on machine learning-based fleet monitoring

Deepti Kunte

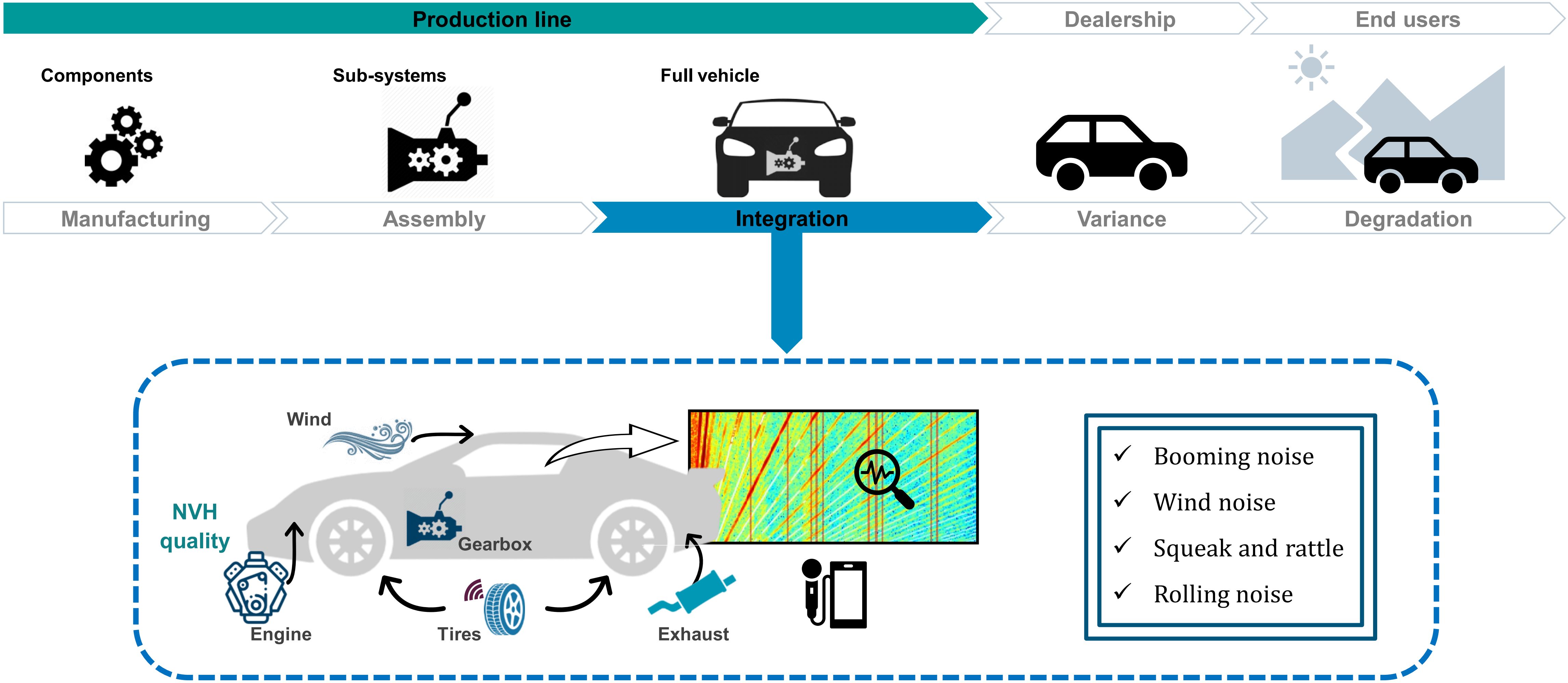
Bram Cornelis

Konstantinos Gryllias

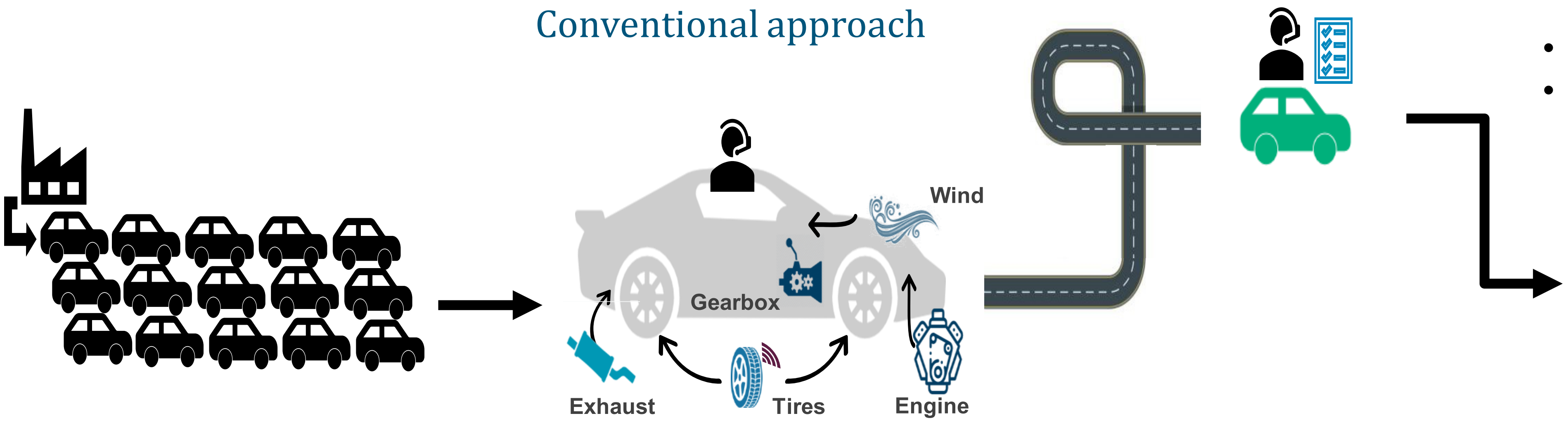
KU LEUVEN SIEMENS

Contents

End-of-line testing and monitoring in fleets



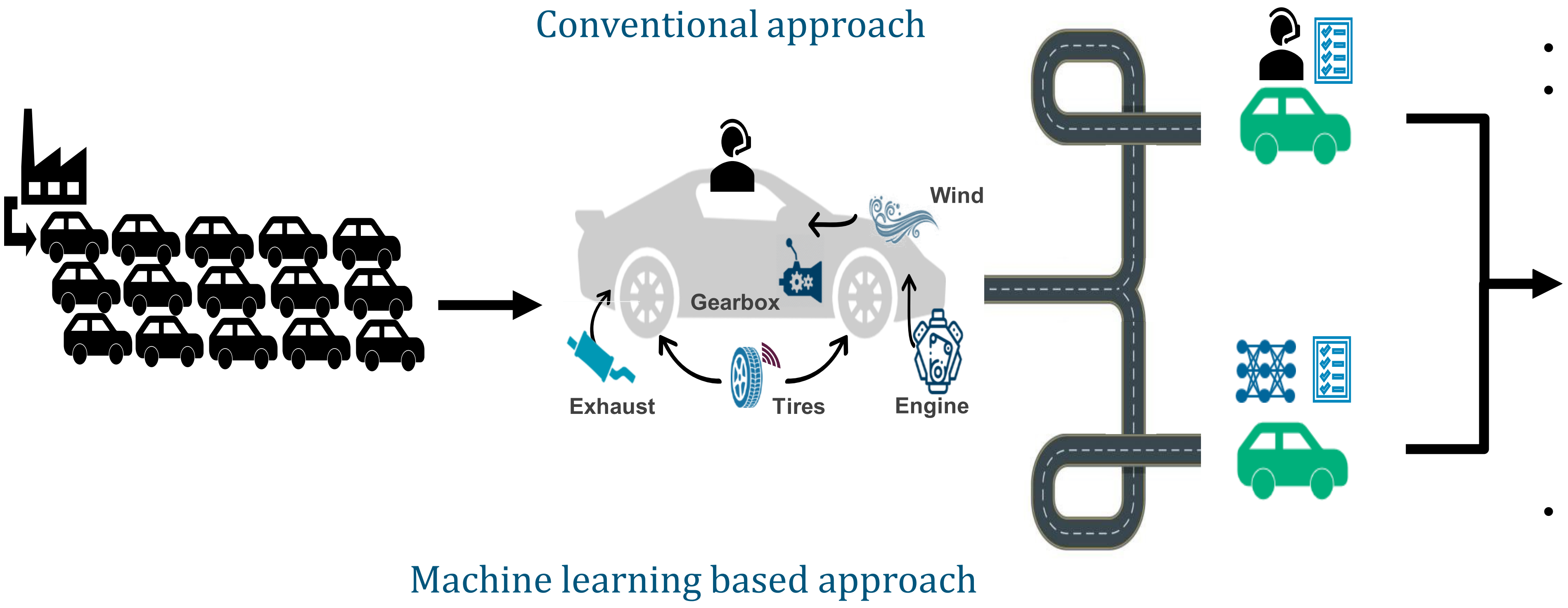
End of line testing



- Reliance on experts
- Non-uniform quality control

KU LEUVEN SIEMENS

End of line testing

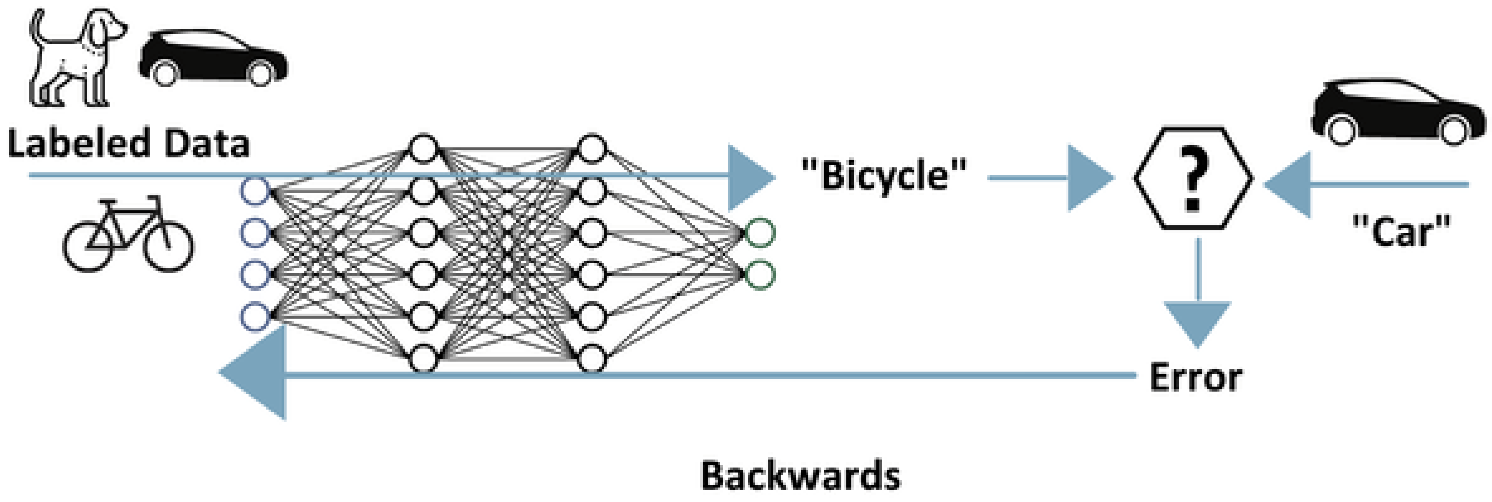


- Reliance on experts
- Non-uniform quality control

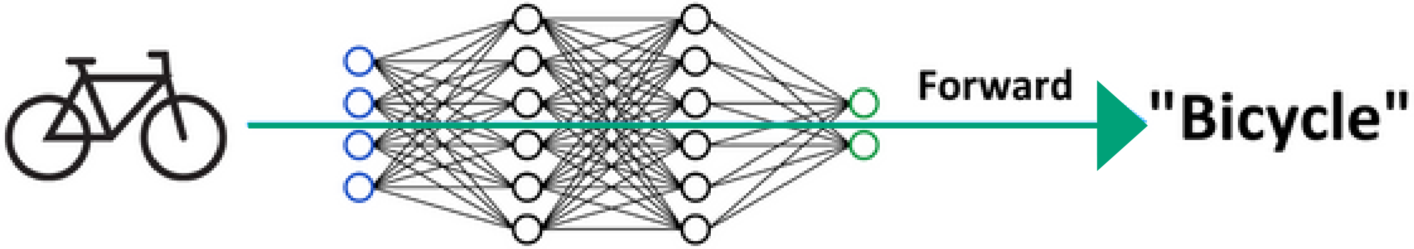
- Need for large datasets

Machine learning stages

Training



Inference



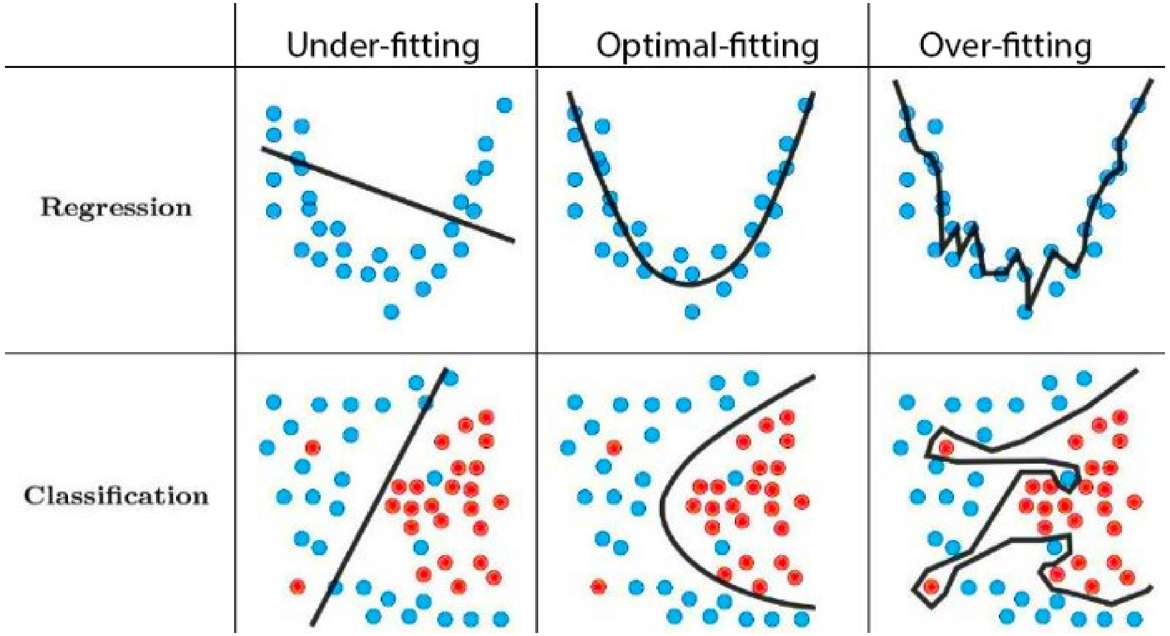
KU LEUVEN SIEMENS

Machine learning stages

Model works well during training but fails on deployment.

Reasons for failure

- Overfitting

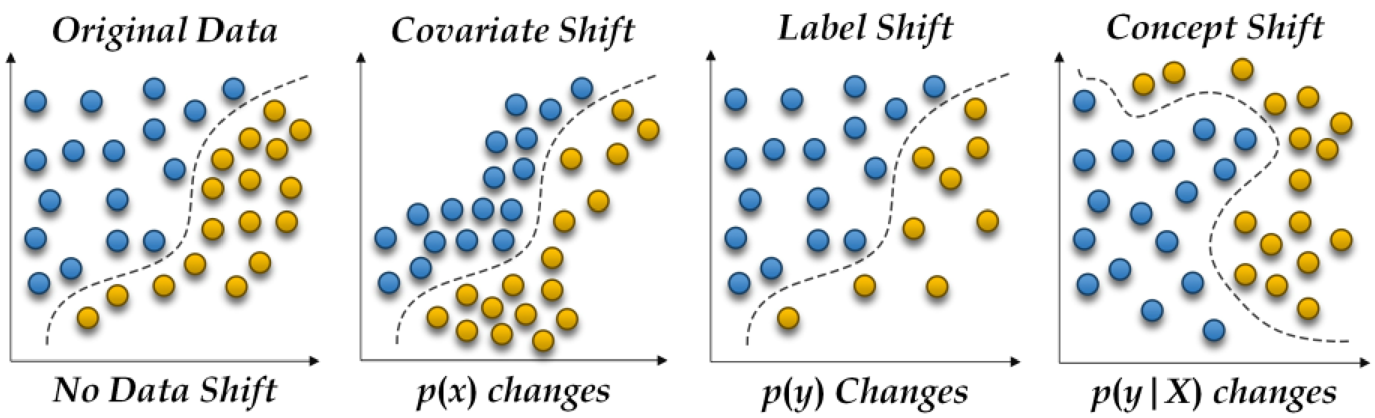


Machine learning stages

Model works well during training but fails on deployment.

Reasons for failure

- Overfitting
- Dataset shift



Machine learning stages

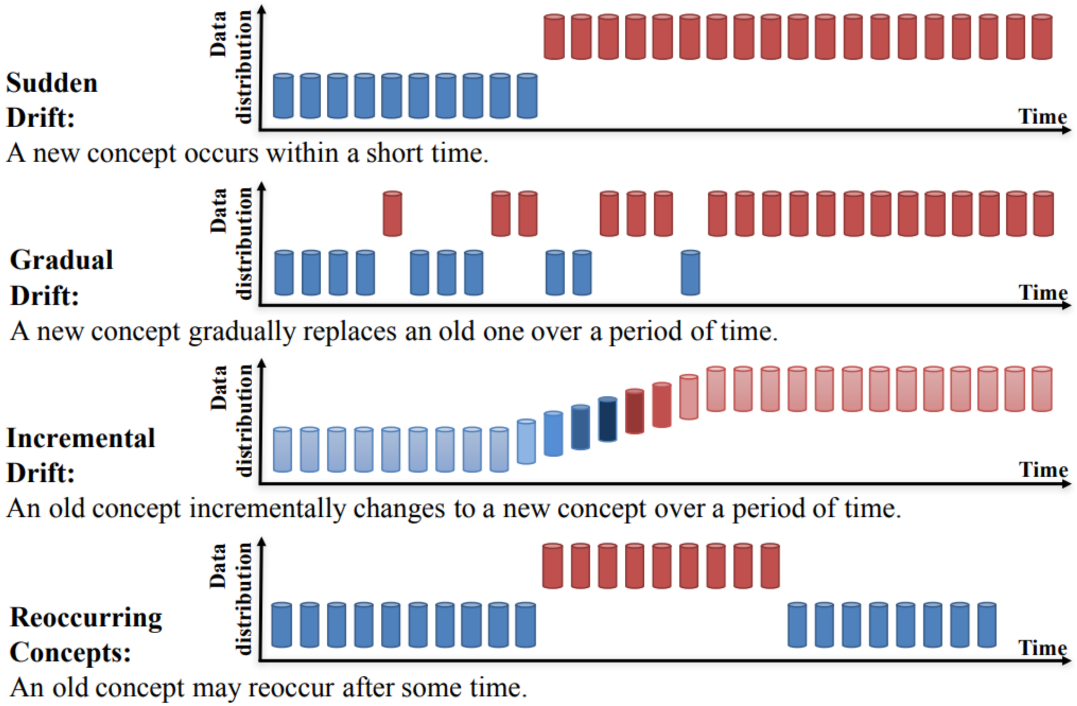
Model works well during training but fails on deployment.

Reasons for failure

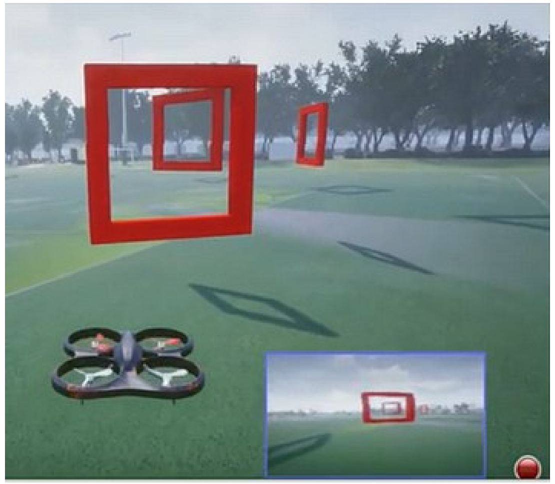
- **Overfitting**: Model memorizes training data and does not generalize to the testing data
- **Dataset shift**: The joint distributions of the training and testing datasets differ and the model is no longer valid on the new data

Examples of dataset shift

The statistical properties of the data that was used to train a machine learning model can change over time. This can cause the model to become less accurate or perform differently than it was designed to.



Simulated data is used to train the models. However, the models will be applied on real life conditions.



Simulation condition

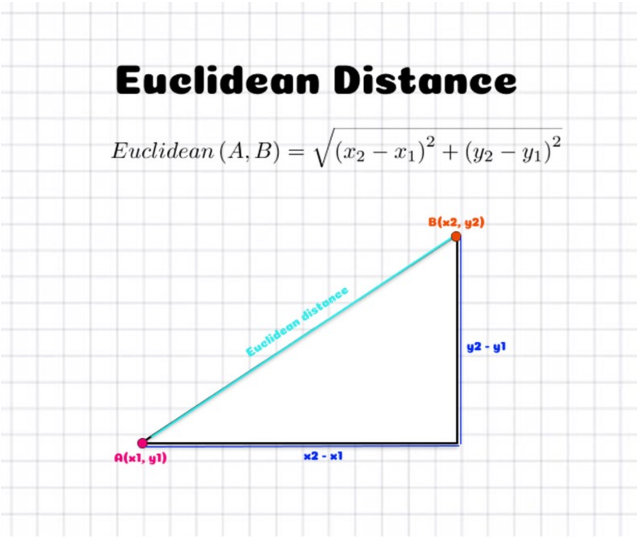
Real life condition

KU LEUVEN SIEMENS

Dataset shift detection and quantification

Discrepancy measures

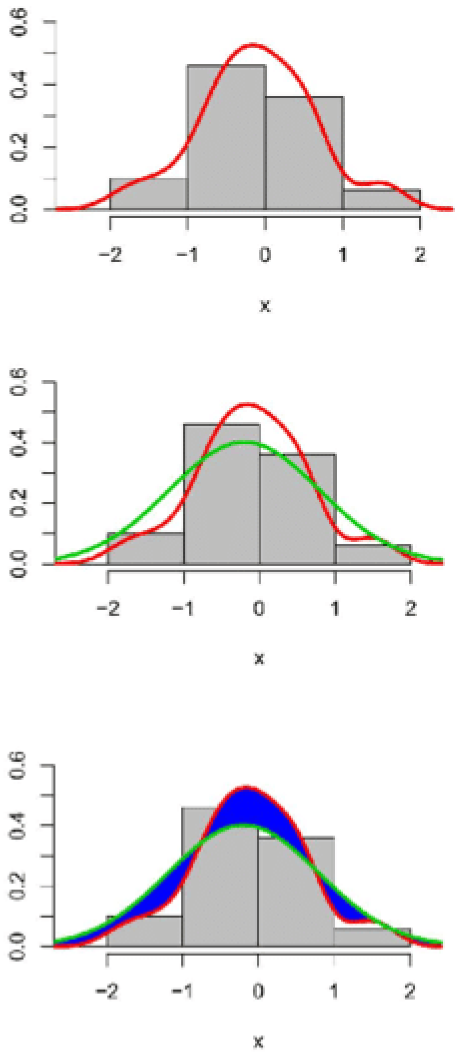
Geometric



KU LEUVEN SIEMENS

Dataset shift detection and quantification

- Discrepancy measures
  - Geometric
  - Information theoretic

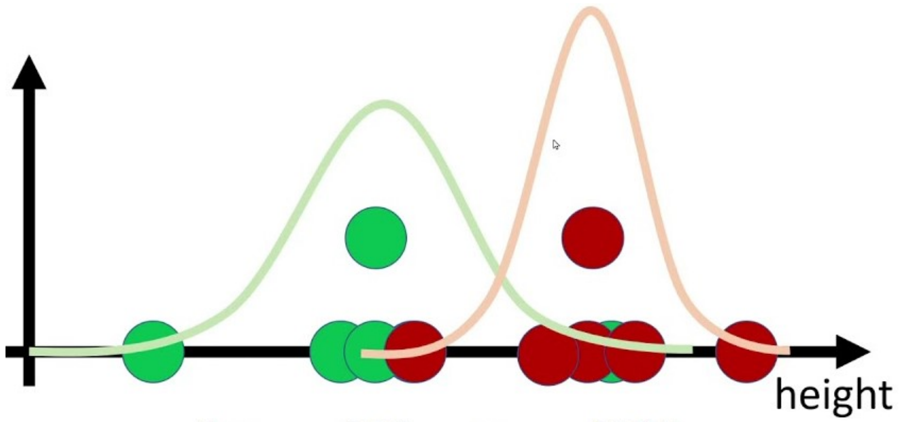


Mis-Specification Analysis of Wiener Degradation Models by Using f-Divergence with Outliers - Fode Zhang, Hon Keung Tony Ng, Yimin Shi

Dataset shift detection and quantification

Discrepancy measures

- Geometric
- Information theoretic
- Higher order



$$\|\mathbb{E}_{X\sim P}[X] - \mathbb{E}_{Y\sim Q}[Y]\|$$

Dataset shift detection and quantification

- Discrepancy measures
  - Geometric
  - Information theoretic
  - Higher order

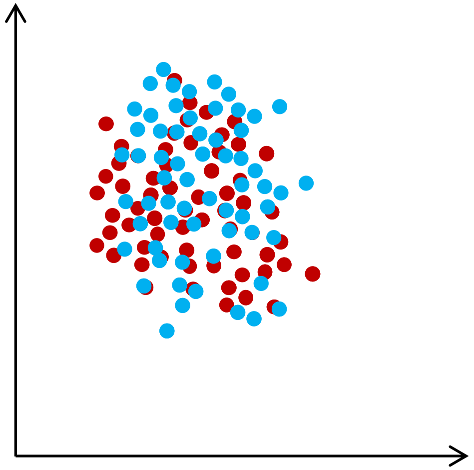
Geometric measures calculate the distance between two vectors in a metric space.

Information-theoretic measures captures the distance between probability distributions.

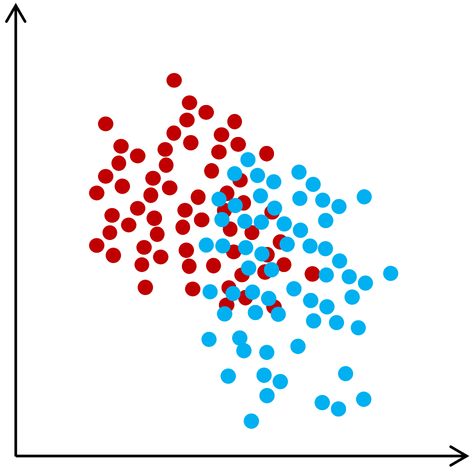
Higher-Order measures consider matching higher order moments of random variables or divergence in a projected space.

Proxy A distance

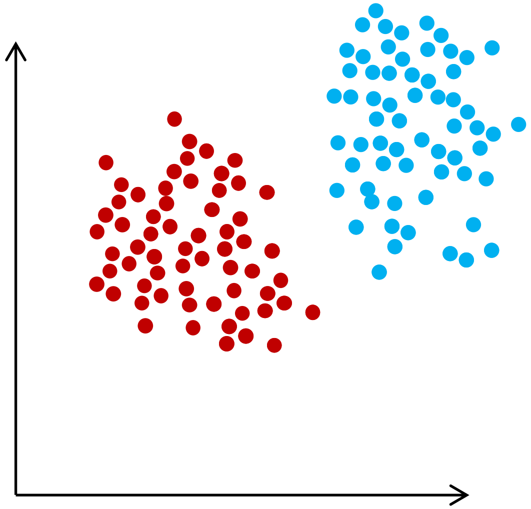
- Domain discrepancy is calculated using a domain classifier
- PAD = 2*(2a-1) where a is the accuracy of the classifier on the test set



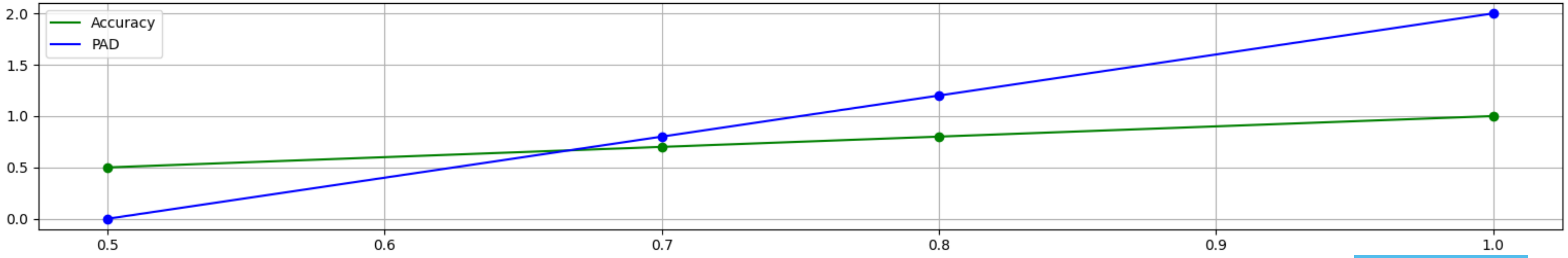
Not separable



Partially separable



Fully separable



Dataset shift simulation

Class distribution shift

- Imbalanced data is the shift caused due to a difference in the proportion of different fault classes in the source and target datasets

Dataset 1	Dataset 2
% Fault	% Fault
50	0
	10
	20
	30
	40
	50
	60
	70
	80
	90
	100

Mixture component shift

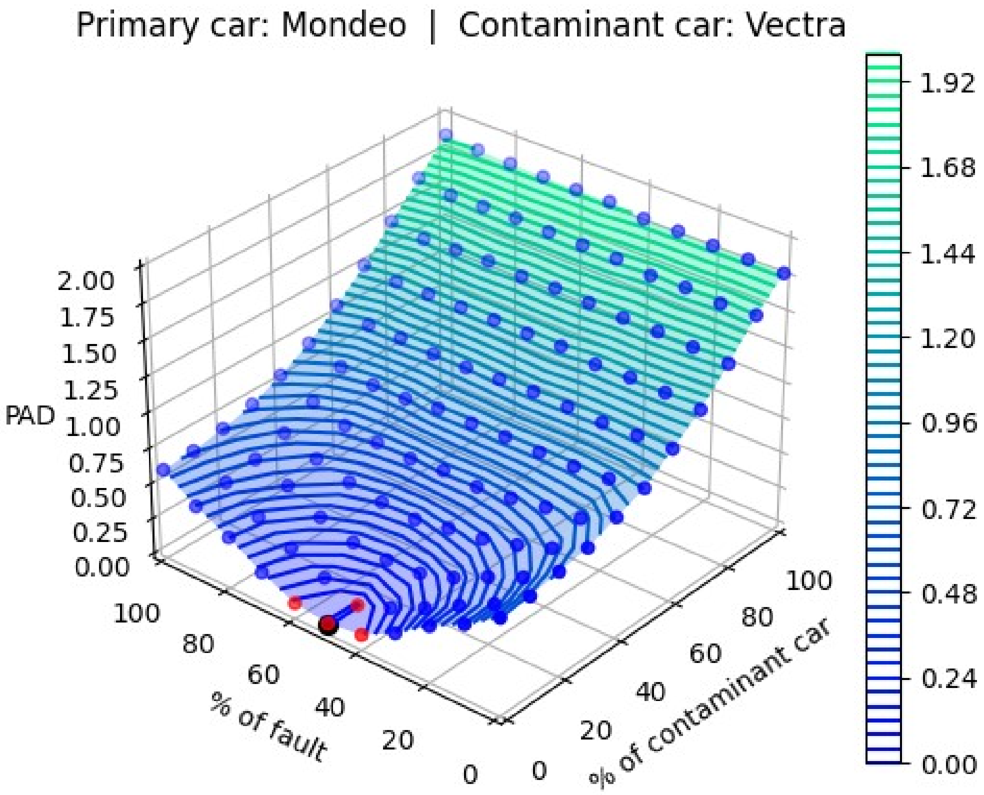
- If the global distribution is made up of data from different sub-populations with varying characteristics, the differences in the proportions of these sub-populations in the two datasets sets leads to a dataset shift called the mixture component shift

Dataset 1	Dataset 2	
% Primary car	% Contaminant car	% Primary car
100	100	0
	90	10
	80	20
	70	30
	60	40
	50	50
	40	60
	30	70
	20	80
	10	90
	0	100

KU LEUVEN

SIEMENS

Simulating both shifts together



KU LEUVEN SIEMENS

Overcoming dataset shift

- Collect and label new dataset

- Build model on new dataset

Overcoming dataset shift

- Collect and label new dataset – Collection of real experimental data is resource intensive
- Build model on new dataset – Building new models for each change can be computationally expensive

How to deal with dataset shift – Transfer learning

- **Transfer learning** is the idea of re-using knowledge learned in one situation for another situation
- A transfer is done **from a source** domain and task **to a target** domain and task
- The **domain** consists of the input feature space X and the marginal probability distribution p(X)
- The **task** is the predictive function learned from training data f

Hadrien Bertrand. Hyper-parameter optimization in deep learning and transfer learning : applications to medical imaging.

Solution categorization

Solution categorization

Transfer knowledge
via the data

Data level

Transfer knowledge by
means of the model

Model level

Zhuang, Fuzhen, et al. "A comprehensive survey on transfer learning." Proceedings of the IEEE 109.1 (2020): 43-76.
https://medium.com/georgian-impact-blog/transfer-learning-part-1-ed0c174ad6e7#071d

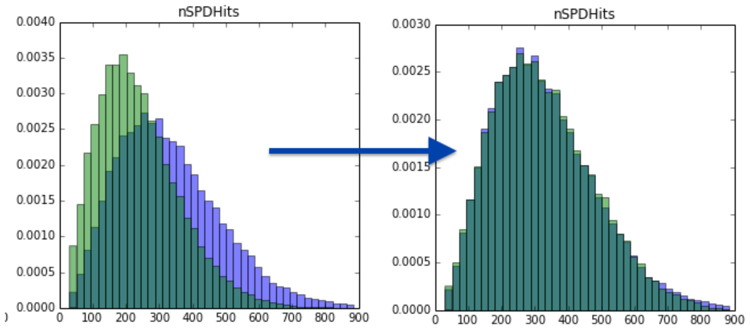
Solution categorization

Solution categorization

Transfer knowledge via the data

Data level

Instance based approaches



Transfer knowledge by means of the model

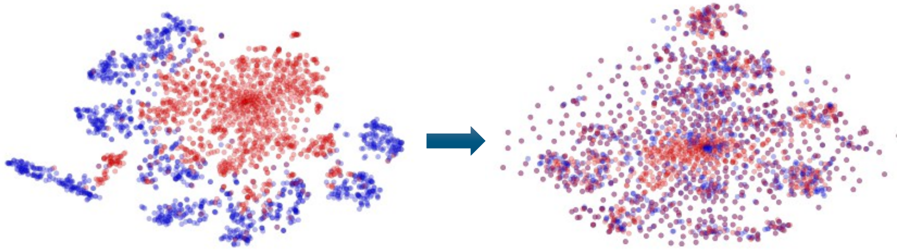
Model level

Zhuang, Fuzhen, et al. "A comprehensive survey on transfer learning." Proceedings of the IEEE 109.1 (2020): 43-76.
https://medium.com/georgian-impact-blog/transfer-learning-part-1-ed0c174ad6e7#071d
https://arogozhnikov.github.io/2015/10/09/gradient-boosted-reweighter.html

Solution categorization

Solution categorization

- Transfer knowledge via the data
  - Data level
    - Instance based approaches
    - Feature based approaches
- Transfer knowledge by means of the model
  - Model level

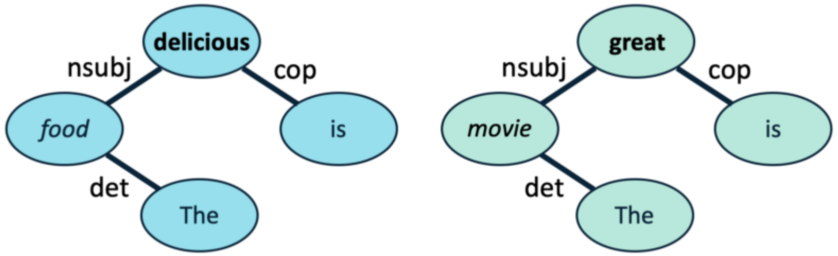


Zhuang, Fuzhen, et al. "A comprehensive survey on transfer learning." Proceedings of the IEEE 109.1 (2020): 43-76.
https://medium.com/georgian-impact-blog/transfer-learning-part-1-ed0c174ad6e7#071d
Ganin, Yaroslav, et al. "Domain-adversarial training of neural networks." The journal of machine learning research 17.1 (2016): 2096-2030.

Solution categorization

Solution categorization

- Data level — Transfer knowledge via the data
  - Instance based approaches
  - Feature based approaches
  - Relation based approaches
- Model level — Transfer knowledge by means of the model



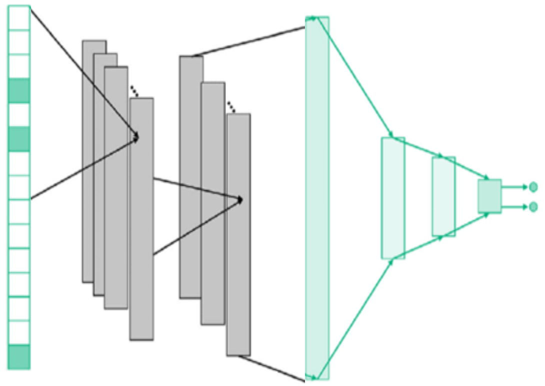
Camera Domain

Movie Domain

Zhuang, Fuzhen, et al. "A comprehensive survey on transfer learning." Proceedings of the IEEE 109.1 (2020): 43-76.
https://medium.com/georgian-impact-blog/transfer-learning-part-1-ed0c174ad6e7#071d

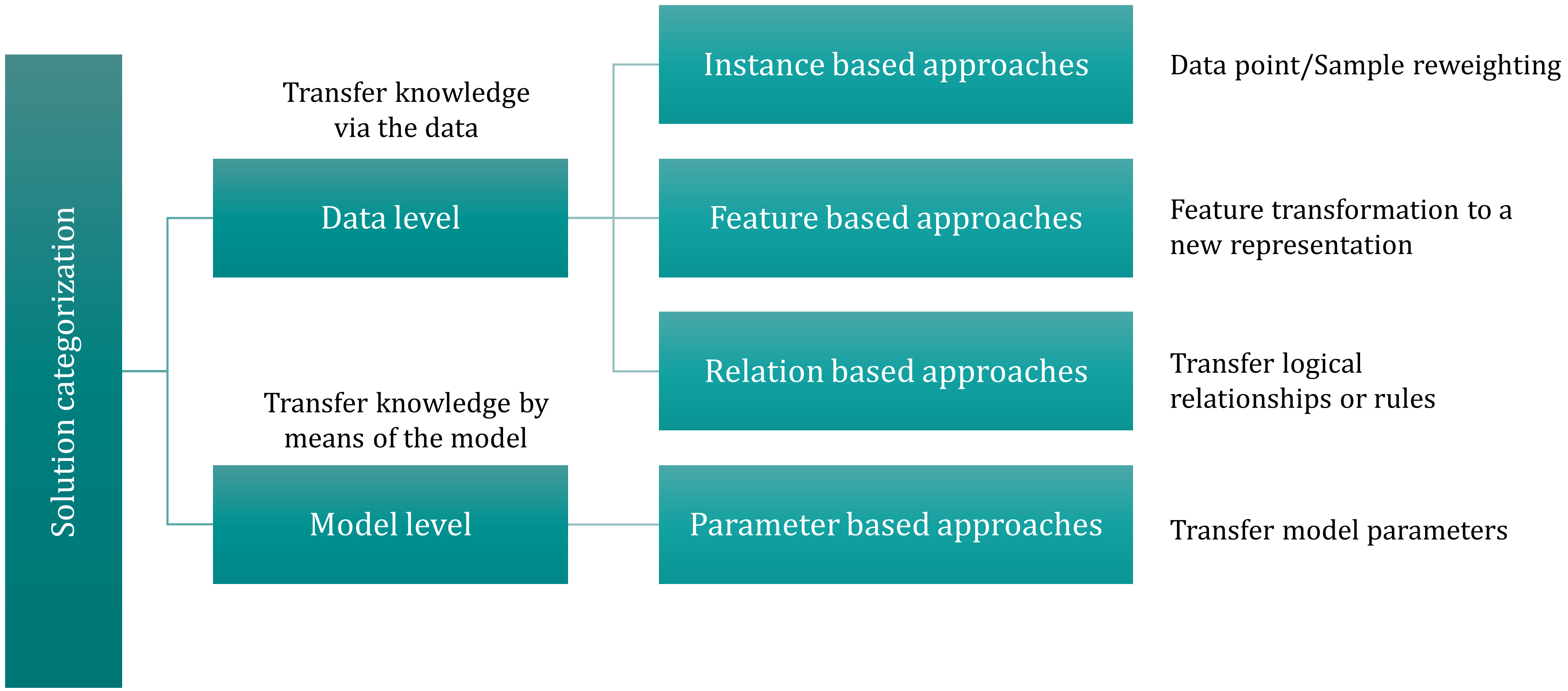
Solution categorization

- Solution categorization
  - Data level (Transfer knowledge via the data)
    - Instance based approaches
    - Feature based approaches
    - Relation based approaches
  - Model level (Transfer knowledge by means of the model)
    - Parameter based approaches



Zhuang, Fuzhen, et al. "A comprehensive survey on transfer learning." Proceedings of the IEEE 109.1 (2020): 43-76.
https://medium.com/georgian-impact-blog/transfer-learning-part-1-ed0c174ad6e7#071d

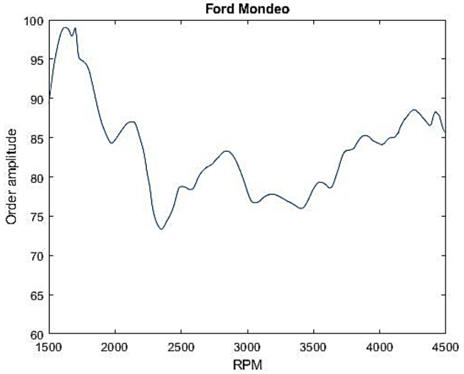
Solution categorization

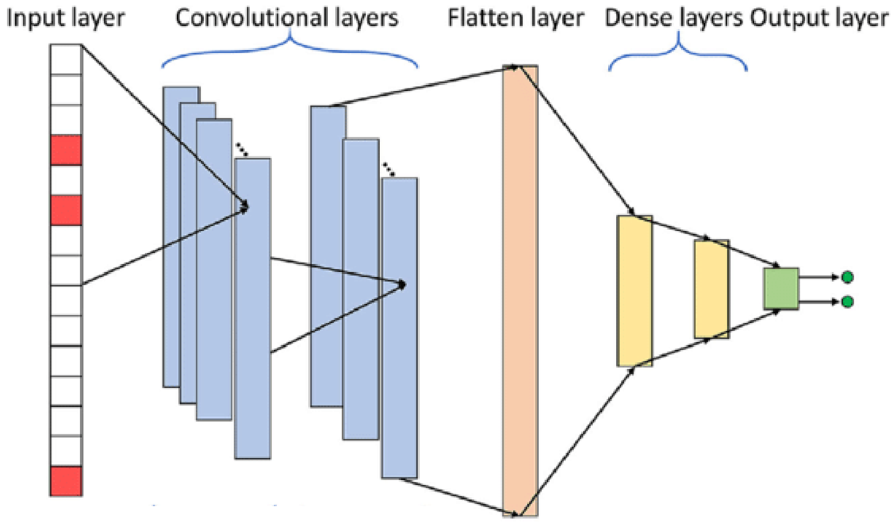


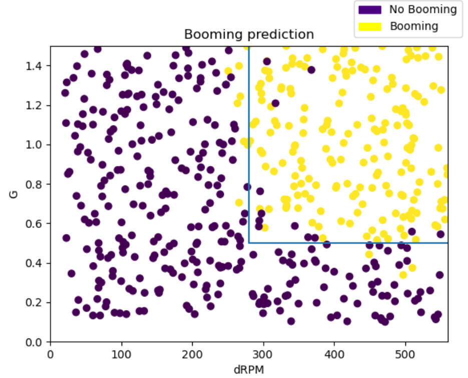
Zhuang, Fuzhen, et al. "A comprehensive survey on transfer learning." Proceedings of the IEEE 109.1 (2020): 43-76.
https://medium.com/georgian-impact-blog/transfer-learning-part-1-ed0c174ad6e7#071d

Booming noise classification

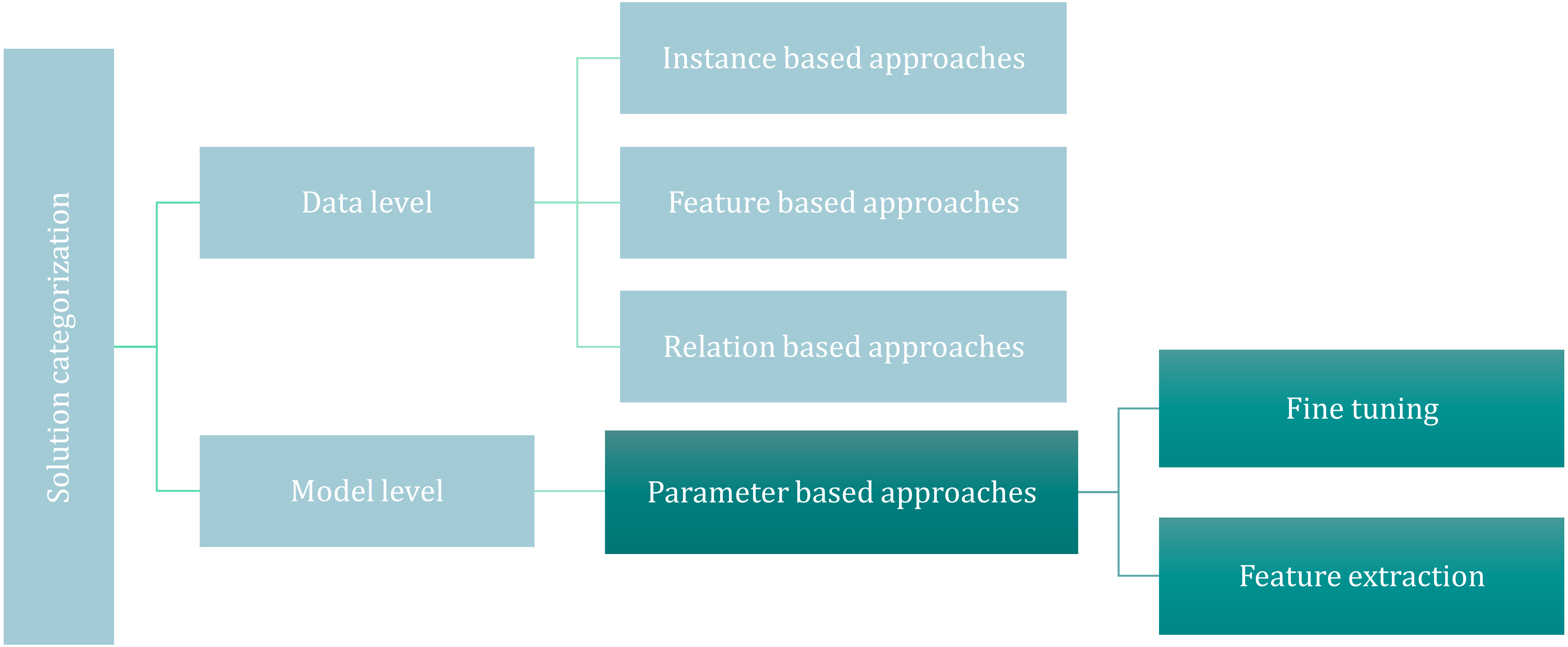
Input profiles → 1D CNN → Booming/ Non-booming







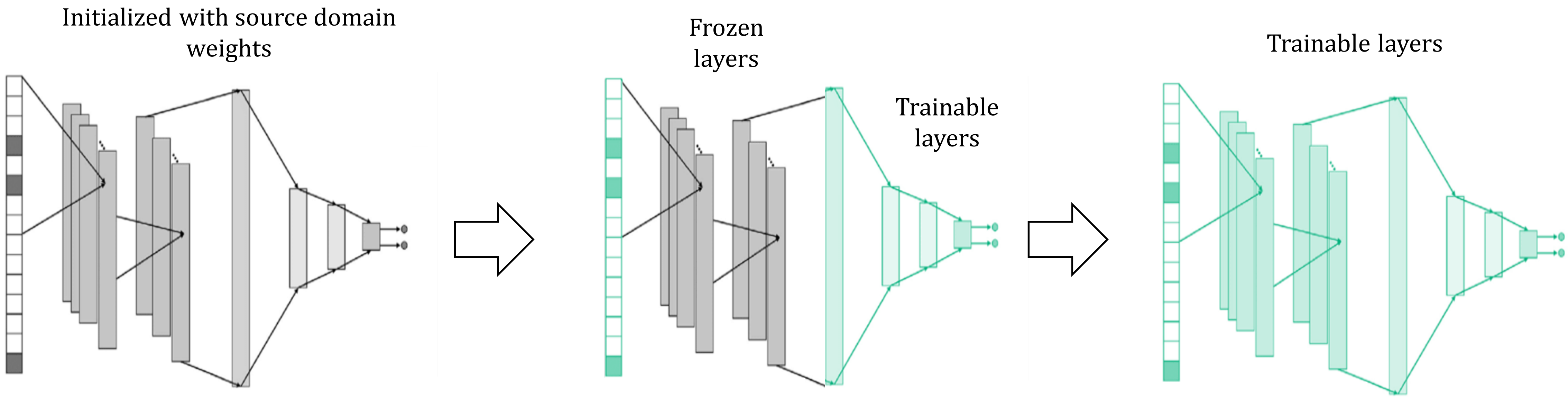
Scarcity of samples



Zhuang, Fuzhen, et al. "A comprehensive survey on transfer learning." Proceedings of the IEEE 109.1 (2020): 43-76.
https://medium.com/georgian-impact-blog/transfer-learning-part-1-ed0c174ad6e7#071d

Fine tuning

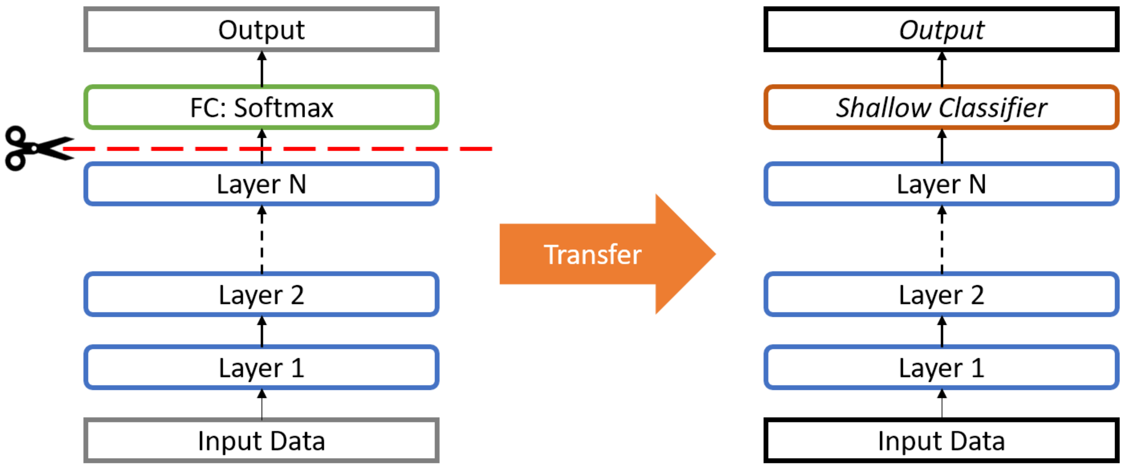
- In this method, a neural network is first trained on the source domain.
- The lower layers which capture more generic features are frozen, while the end layers are further trained on the target domain
- As a last step the entire model can be further trained on the target dataset



https://livebook.manning.com/book/deep-learning-with-javascript/chapter-5/v-6/114

Feature extraction

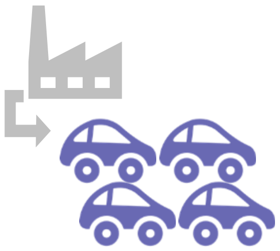
- Similar to fine-tuning, a neural network is first trained on the source domain.
- The output of the lower layers is then used as input to a completely different model which is trained on the target domain from scratch.
- This new model need not be a neural network.



https://www.oreilly.com/library/view/hands-on-transfer-learning/9781788831307/d94586c6-1c46-4794-aded-22442a4f81d8.xhtml

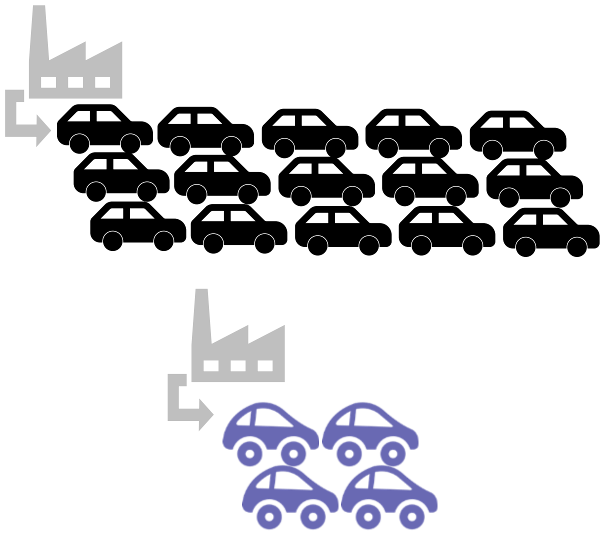
Results

Small target dataset



Test accuracy: 78.80 %

Large source dataset
+ Small target dataset

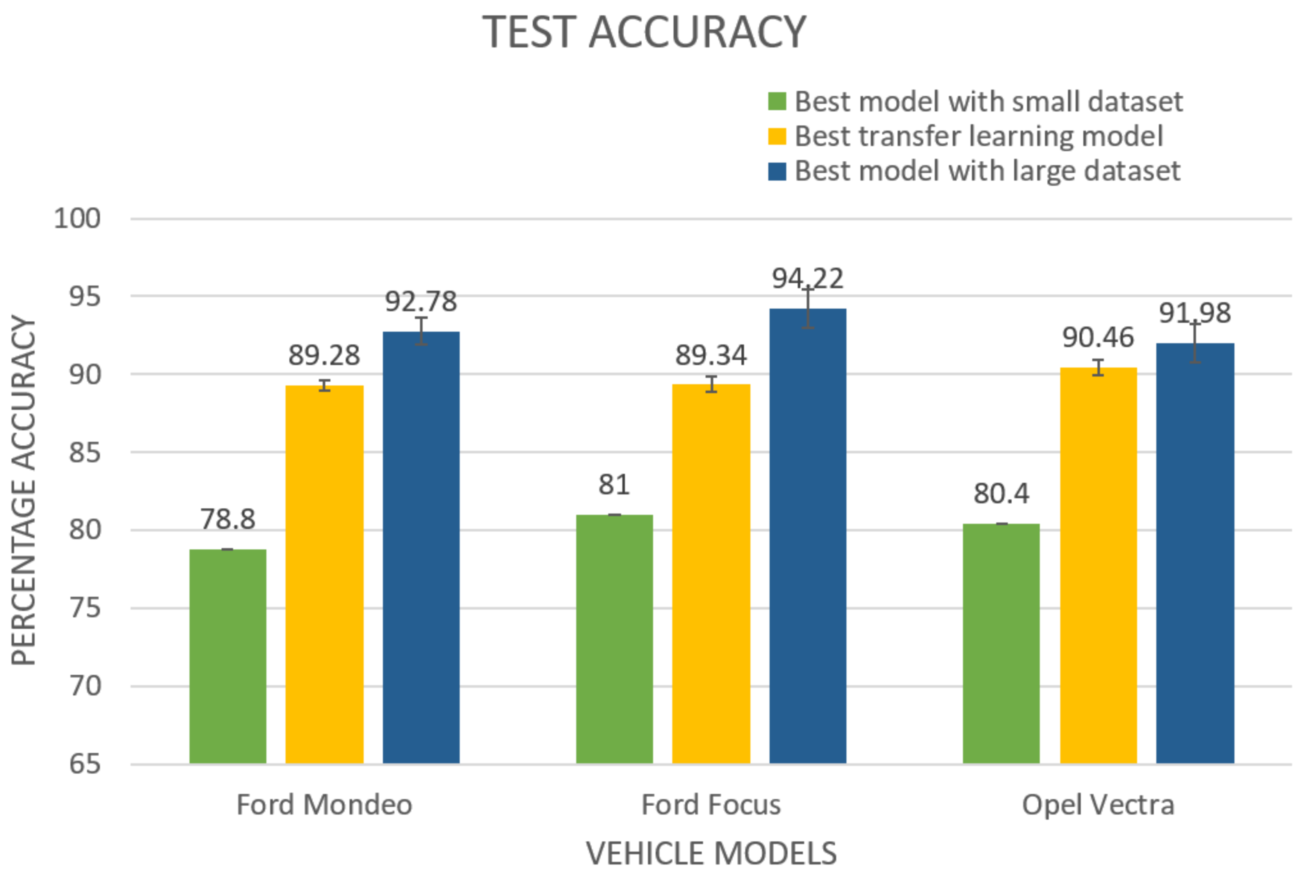


Test accuracy: 89.28 %

Large target dataset

Test accuracy: 92.78 %

Results



KU LEUVEN SIEMENS

Data scarcity

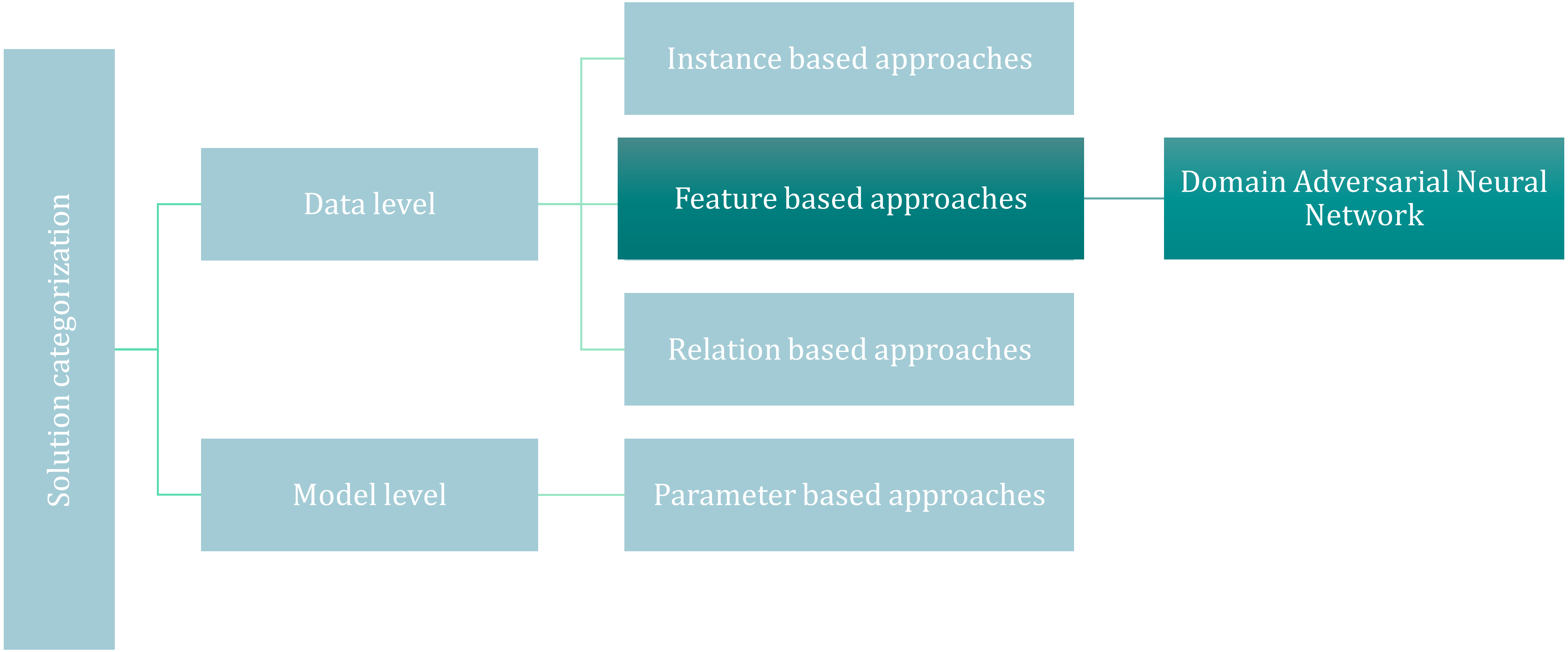
- Scarcity of samples
  - Access to **few sound samples** from the target domain
  - Target domain is **labelled**

Dataset	Input features used	Training + Validation	Testing	Labelled
Source	2nd order profile	1000 + 100	500	Yes
Target	2nd order profile	**100 + 100**	500	**Yes**

- Scarcity of labels
  - Access to **sufficient sound samples** from the target domain
  - Target domain is **unlabelled**

Dataset	Input features used	Training + Validation	Testing	Labelled
Source	2nd order, loudness, sharpness profiles	1000 + 250	1000	Yes
Target	2nd order, loudness, sharpness profiles	**1000 + 250**	1000	**No**

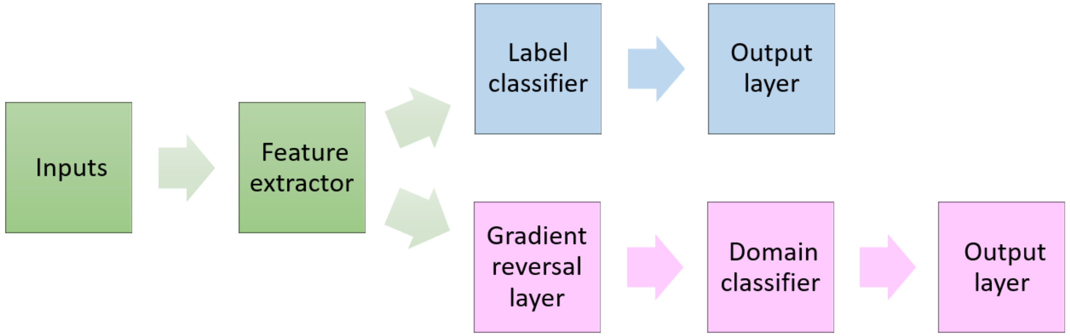
Scarcity of labels



Zhuang, Fuzhen, et al. "A comprehensive survey on transfer learning." Proceedings of the IEEE 109.1 (2020): 43-76.
https://medium.com/georgian-impact-blog/transfer-learning-part-1-ed0c174ad6e7#071d

Domain Adversarial Neural Network (DANN)

- Feature extractor extracts domain invariant and discriminative feature space
- In the forward propagation, the features are sent to:
  1. A binary classifier *Gd* to classify whether they come from the source or the target domain with domain label *d*.
  2. For the source domain data, the features are simultaneously sent to the label classifier *Gy* to predict class label *y*
- DANN seeks to minimize the source classification loss for the discriminativeness while maximizing the domain classification loss for the domain-invariance.
- The Gradient Reversal Layer (GRL) serves for an identity transformation in the forward propagation, while the downstream gradients will change the sign passing through the GRL during backpropagation.



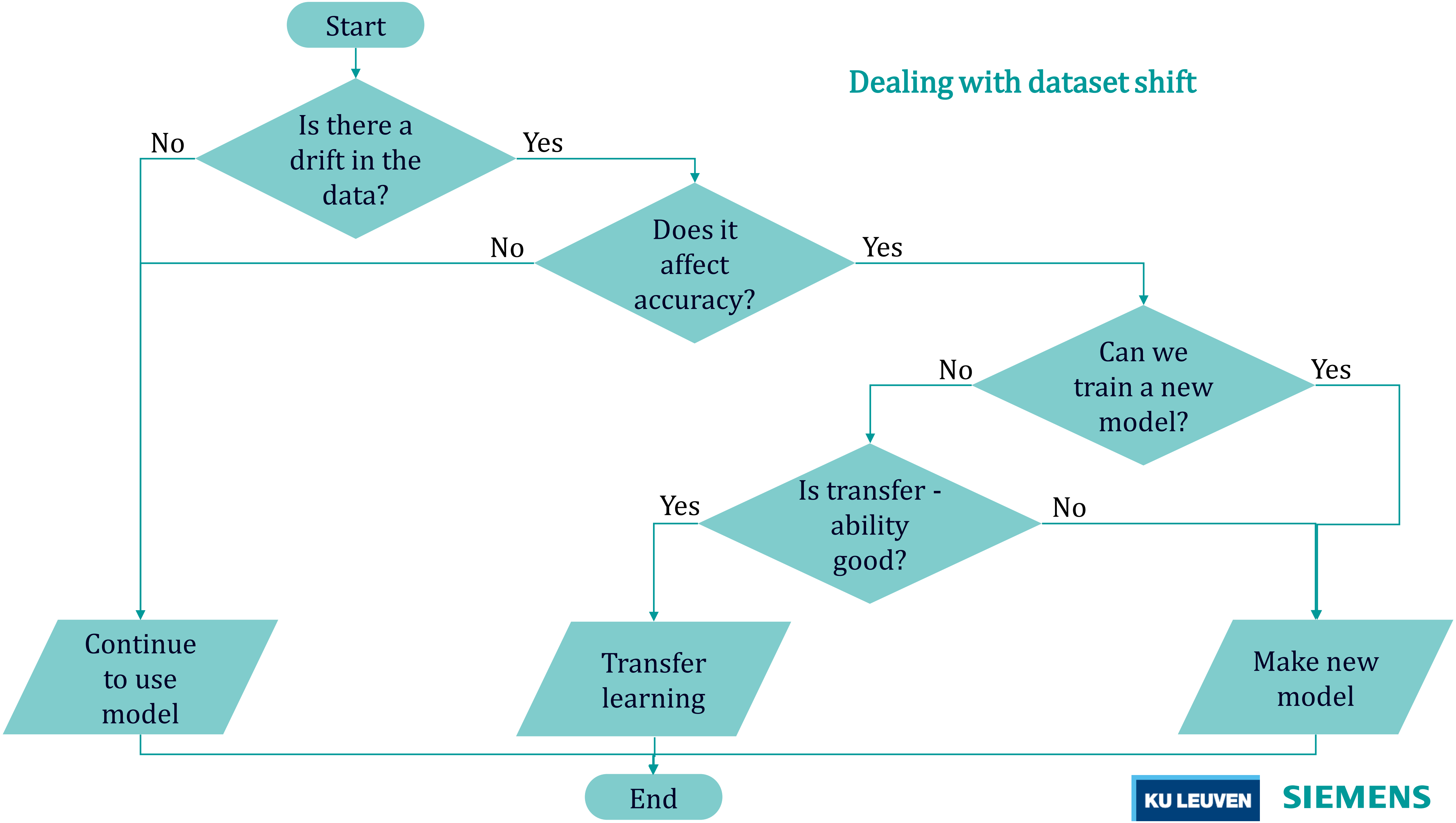
$$\mathcal{L}(\theta_f,\theta_y,\theta_d)=\frac{1}{n}\sum_{i=1}^{n}\mathcal{L}_y^i(\theta_f,\theta_y)-\lambda\Big(\frac{1}{n}\sum_{i=1}^{n}\mathcal{L}_d^i(\theta_f,\theta_d)+\frac{1}{n'}\sum_{i=1}^{n'}\mathcal{L}_d^i(\theta_f,\theta_d)\Big)$$

Ganin, Yaroslav, et al. "Domain-adversarial training of neural networks."
The journal of machine learning research 17.1 (2016): 2096-2030.

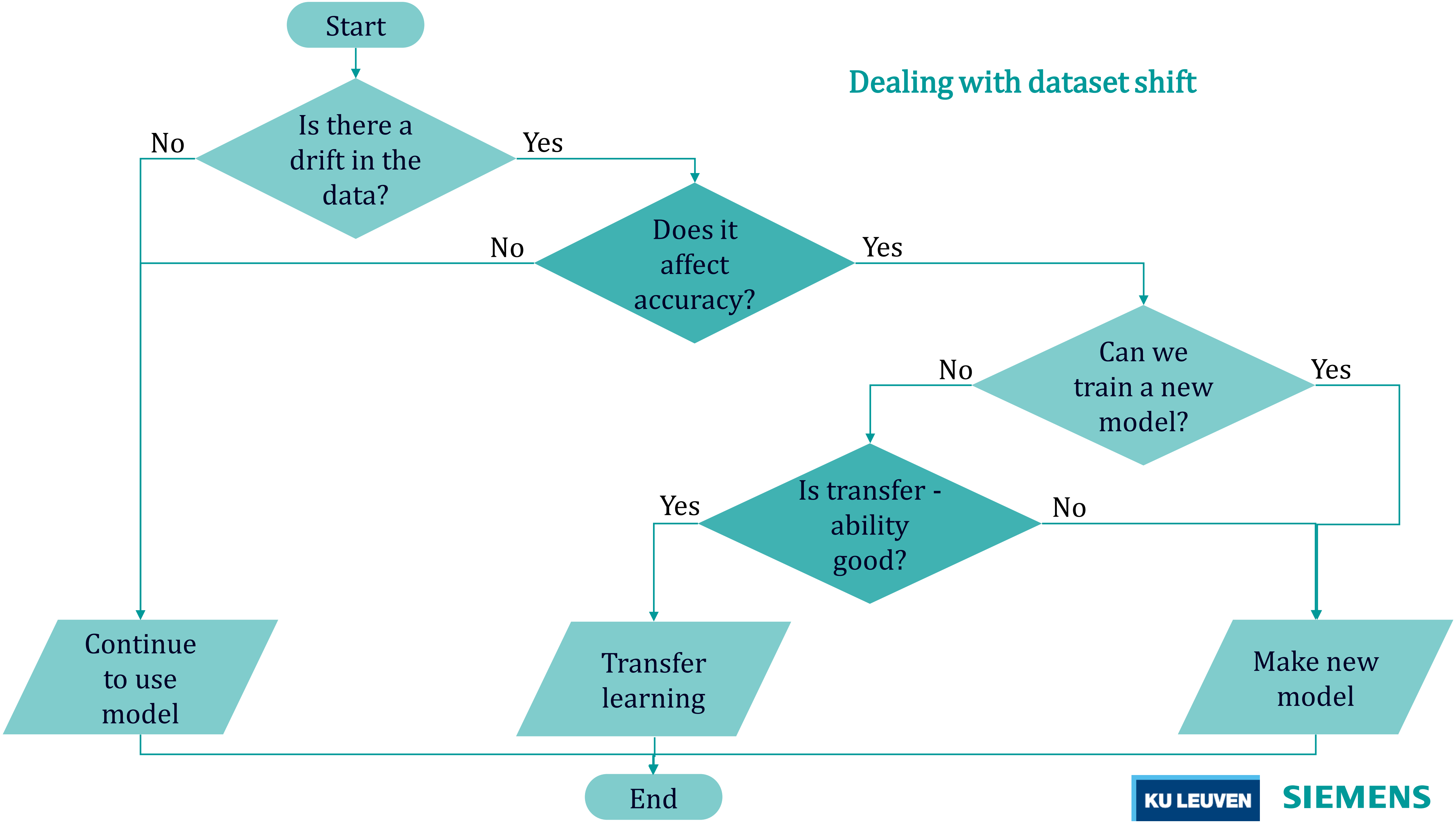
Domain Adversarial Neural Network Results

Source	Target	**Without trend removal** Percentage accuracy (Mean ± standard deviation)			**After trend removal** Percentage accuracy (Mean ± standard deviation)		
		Without DANN	With DANN	Fully labelled dataset	Without DANN	With DANN	Fully labelled dataset
Mondeo	Focus	78.9 ± 2.7	83.1 ± 1.5	93.8 ± 0.1	91.4 ± 0.8	91.5 ± 0.3	93.8 ± 0.7
Vectra		83.5 ± 3.6	85.2 ± 2.5		86.1 ± 1.1	88.0 ± 1.1	
Focus	Mondeo	71.9 ± 14.3	86.6 ± 0.6	93.2 ± 0.8	89.9 ± 1.3	91.5 ± 0.6	94.0 ± 0.7
Vectra		70.9 ± 25.6	86.6 ± 0.8		88.9 ± 1.2	90.6 ± 0.3	
Focus	Vectra	27.1 ± 3.9	82.3 ± 2.8	92.8 ± 0.4	85.3 ± 1.0	88.1 ± 1.2	93.8 ± 0.7
Mondeo		35.9 ± 10.3	83.4 ± 1.0		87.1 ± 0.2	88.5 ± 0.5	

Dealing with dataset shift



Dealing with dataset shift



KU LEUVEN SIEMENS

Future work

- Assessing consequences of shift
- Estimation of transferability

Thank You!

We gratefully acknowledge the European Commission for
its support of the Marie Sklodowska Curie program
through the H2020 ETN MOIRA project (GA 955681)

LMSD Mecha(tro)nic System Dynamics KU LEUVEN SIEMENS