The eXPose Approach to Crosslier Detection

António Pereira Barata[1,2], Frank W. Takes[1], H. Jaap van den Herick[3], and Cor J. Veenman[1,4]

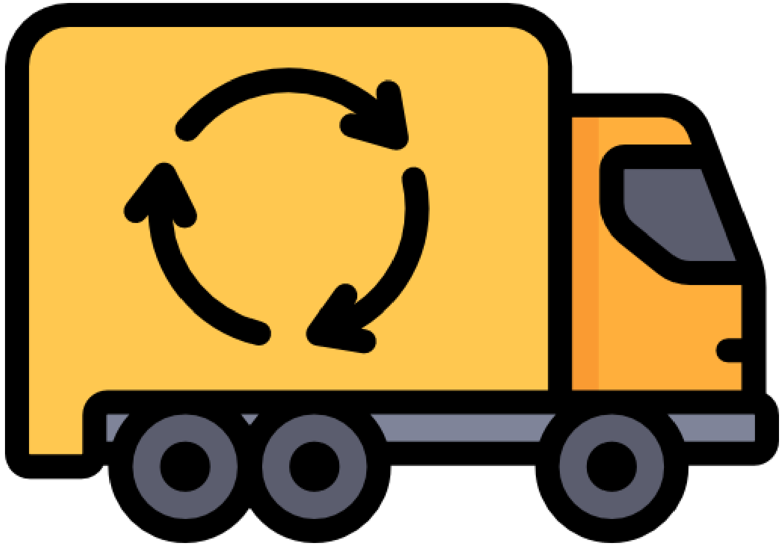
[1]LIACS, Leiden Univeristy, the Netherlands
[2]ILT, Ministry of Infrastructure and Water Management, the Netherlands
[3]LCDS, Leiden University, the Netherlands
[4]Data Science Department, TNO, the Netherlands

25th International Conference on Pattern Recognition | January 2021

Universiteit Leiden
The Netherlands

Inspectie Leefomgeving en Transport
Ministerie van Infrastructuur en Waterstaat

1. Introduction



Waste transportation

- List of Waste categorisation
- Health and environmental risks

Human Environment and Transport Inspectorate (ILT)

- Permit is compliant
- Permit is requires investigation

Companies

- Different waste costs
- Incentive for mislabelling

Inspectors

- Too many permits for manual inspection
- Automatic methods are required

2. Problem

Crosslier

- Category-labelled dataset
- Label and feature swaps
- Crosslyingness

Formally

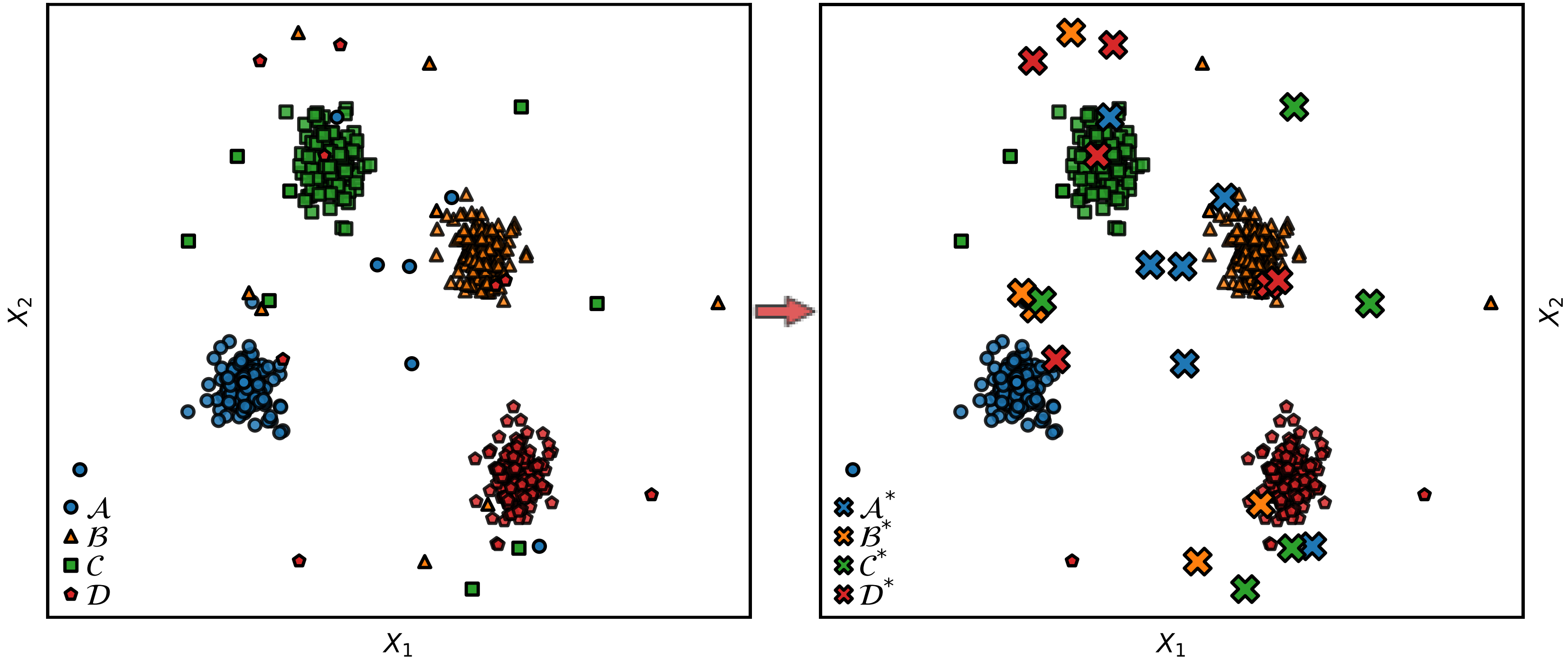
- Distribution **D** random variables $(X, Z) \in \mathbf{X} \times \mathbf{Z}$
- $\mathbf{X} \subseteq \Re^{m}$, $\mathbf{Z} = \{z_1, z_2, \ldots, z_q\}$, q categories
- Samples $(x_1, z_1), \ldots, (x_n, z_n)$

Goal:

- Find $f_z(x)$ for each $z \in \mathbf{Z}$
- Output crosslier score
- $x_i \in X$, $z_i = z$

Domain-agnostic

- Other fields



3. Related Work

Supervised & Semi-supervised

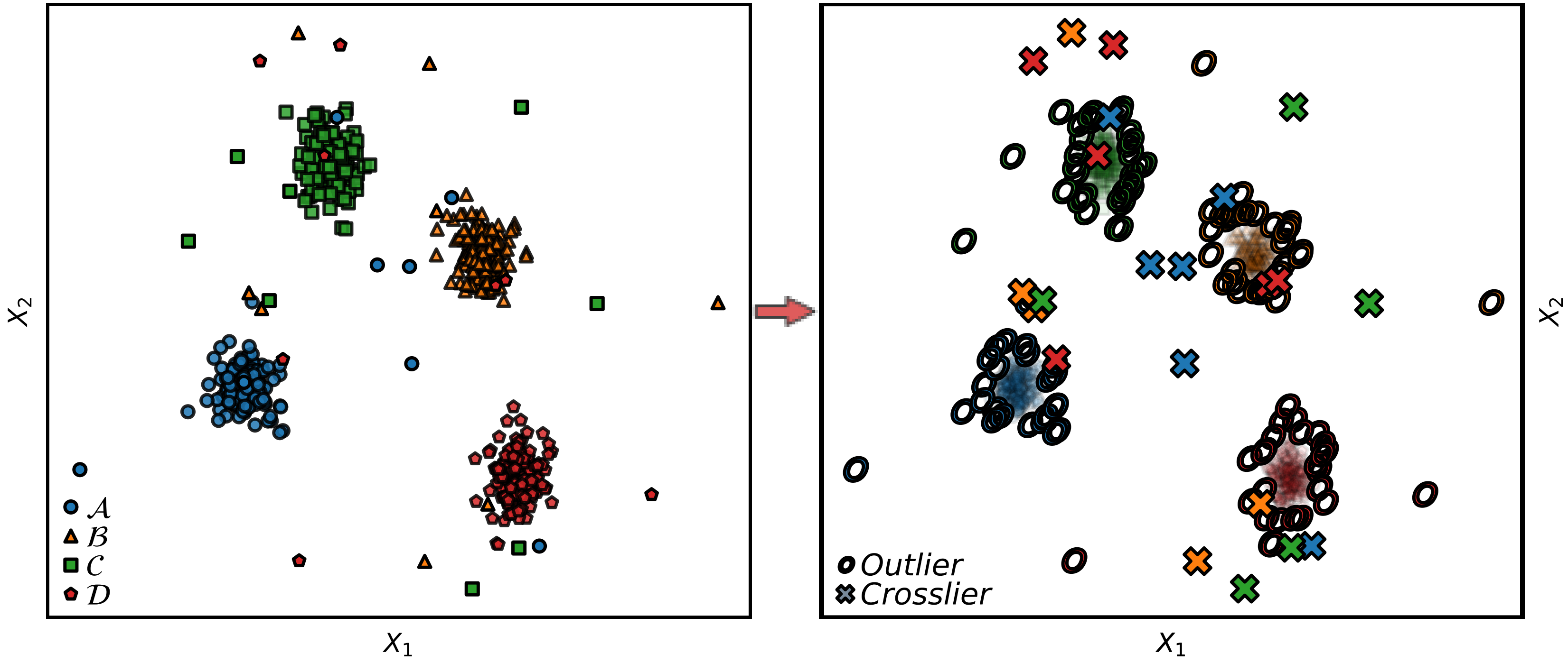
- Non-compliance labelling
- Prediction of labels
- *Insurance fraud detection*

Unsupervised

- Anomaly/Outlier detection
- No labels required
- *Traffic anomaly flagging*

Issues

- (Semi-)Supervised
  - No labels available
- Unsupervised
  - Scale-dependant outputs
  - Curse of dimensionality
  - Not specific enough
    - Example: Isolation Forest



4. The eXPose Approach

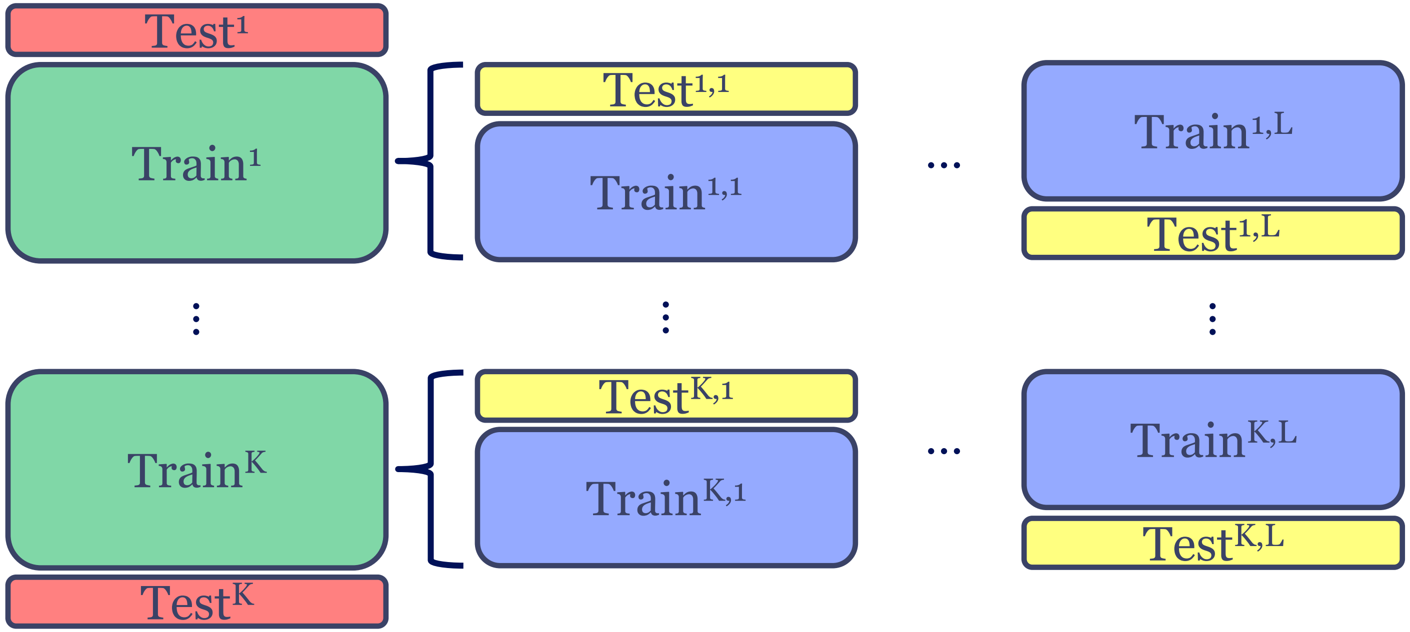
Characteristics

- Data-driven
- Learning functions
- Category-labelled dataset
  - All samples get a crosslier score

In a nutshell

- Supervised learning
- Per category label
- Several learners used
- Scoring on left-out part
  - Cross-validation fashion
    - Calibration (Platt scaling)

$f_z(x) = -\log_2 P(Z=z \mid X=x)$



Inner folds

- Calibration

Outer folds

- Posteriors ⟶ Scores

5. Experiments

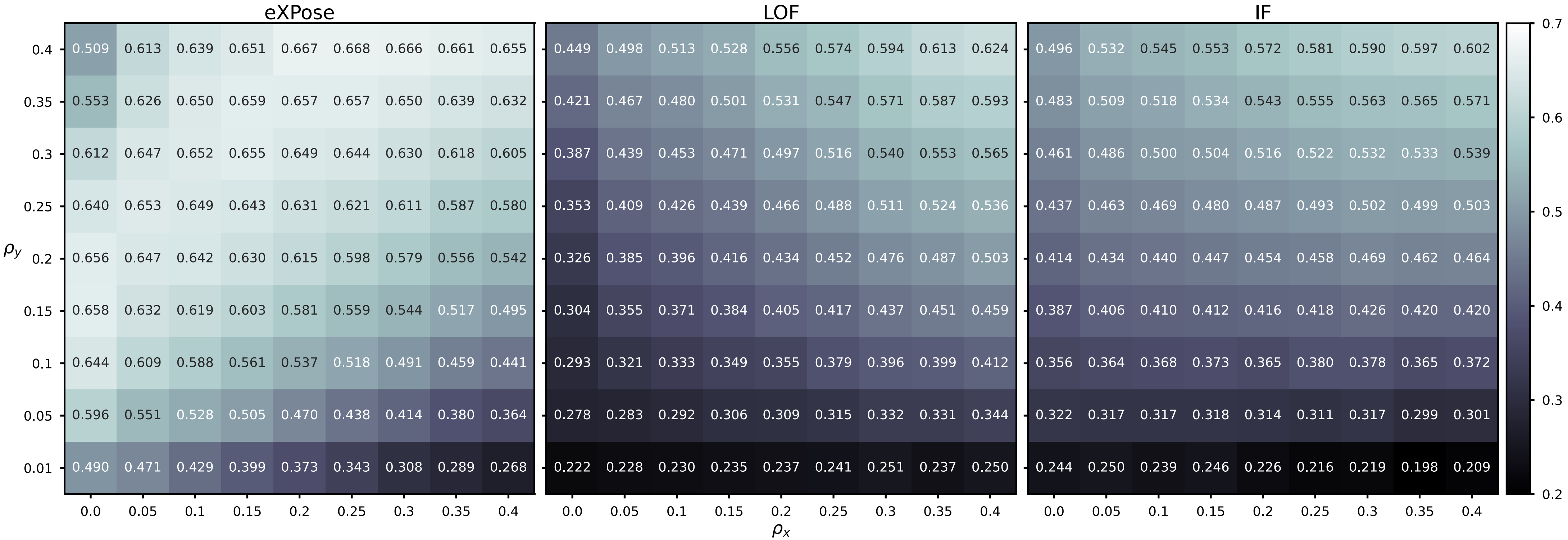
Bechmark

- Comparison
  - Local Outlier Factor (LOF)
  - Isolation Forest (IF)
  - eXPose (XGBoost surrogate model)
- 20 classification datasets
  - *openML.org*
- Synthetic crossliers
  - Label swap $\boldsymbol{\rho}_{\mathbf{y}} \in \{.01, .05, .1, .15, .2, .25, .3, .35, .4\}$
  - Feature replacement $\boldsymbol{\rho}_{\mathbf{x}} \in \{.00, .05, .1, .15, .2, .25, .3, .35, .4\}$
    - Sampled from $\mathbf{D}_{(X \mid Z)}$ estimated as either Normal or Multinomial
  - 10 random initialisation per configuration pair
- Average precision performance

Waste transportation

- 20 waste categories
- Learners
  - Logitic regression (LR)
  - XGBoost (XGB)
- CV model selection by ROC AUC
- Crosslier diagram
  - Domain application

6.1 Results - Benchmark



Cell

- AP mean value
  - Across all datasets
    - For all initialisations

Vertical axis

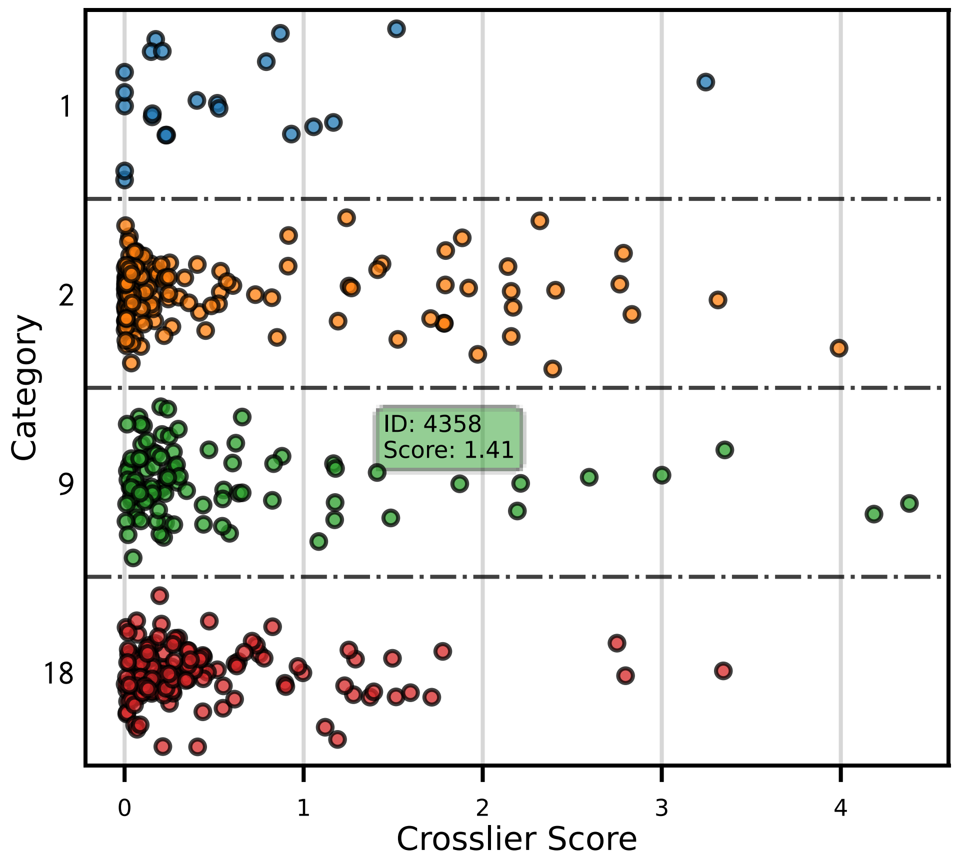
- Proportion of labels swapped

Horizontal axis

- Proportion of features replaced

6.2 Results - Application

Category	LR	XGB
1	0.983 ± 0.008	$\mathbf{0.985 \pm 0.010}$
2	0.868 ± 0.044	$\mathbf{0.919 \pm 0.037}$
3	0.868 ± 0.020	$\mathbf{0.908 \pm 0.027}$
—	—	—
5	0.672 ± 0.092	$\mathbf{0.755 \pm 0.082}$
6	0.740 ± 0.038	$\mathbf{0.794 \pm 0.037}$
7	0.776 ± 0.016	$\mathbf{0.821 \pm 0.015}$
8	0.798 ± 0.026	$\mathbf{0.856 \pm 0.025}$
9	0.867 ± 0.047	$\mathbf{0.915 \pm 0.047}$
10	0.737 ± 0.032	$\mathbf{0.788 \pm 0.035}$
11	0.815 ± 0.021	$\mathbf{0.896 \pm 0.016}$
12	0.860 ± 0.032	$\mathbf{0.897 \pm 0.031}$
13	0.609 ± 0.063	$\mathbf{0.720 \pm 0.062}$
14	0.776 ± 0.034	$\mathbf{0.817 \pm 0.024}$
15	0.841 ± 0.019	$\mathbf{0.883 \pm 0.016}$
16	0.695 ± 0.016	$\mathbf{0.753 \pm 0.019}$
17	0.845 ± 0.023	$\mathbf{0.889 \pm 0.022}$
18	0.894 ± 0.015	$\mathbf{0.921 \pm 0.015}$
19	0.806 ± 0.014	$\mathbf{0.851 \pm 0.013}$
20	0.719 ± 0.024	$\mathbf{0.779 \pm 0.027}$



7. Discussion

Method

- eXPose > outlier methods
  - Expected by definition
    - Specific target
    - Feature values
    - Category-wise
- Empirical density estimation
  - High dimensionality
    - Robust learners
    - Feature selection
    - Regularisation

Implementation

- Supervised category-modelling
- Calibrated scores
  - Reliable output
    - Low AUC: scores of 1
    - eXPose will not *expose*
    - Important in sensitive domains

Application

- ILT Inspectors
  - Crosslier diagram
    - High scoring samples
    - Further investigation
    - Valuable tool **vs** spreadsheet

8. Conclusion

Achievements

- Crosslier
- eXPose approach
- Crosslier diagram

Suitability

- Superior to outlier detection
- Domain-insensitive

Waste category

- Value insights
  - Targetting fraudulent permits
  - Crosslier: company misconduct
- High AUC models
  - Reliable flagging
  - Adequate data collection
  - Compliance checking

Future work

- Further cooperation
  - Feedback
  - Labelling
    - Supervised
- Other domains

Thank you

Universiteit Leiden
The Netherlands

Inspectie Leefomgeving en Transport
Ministerie van Infrastructuur en Waterstaat