

Automated detection of malicious .nl-registrations

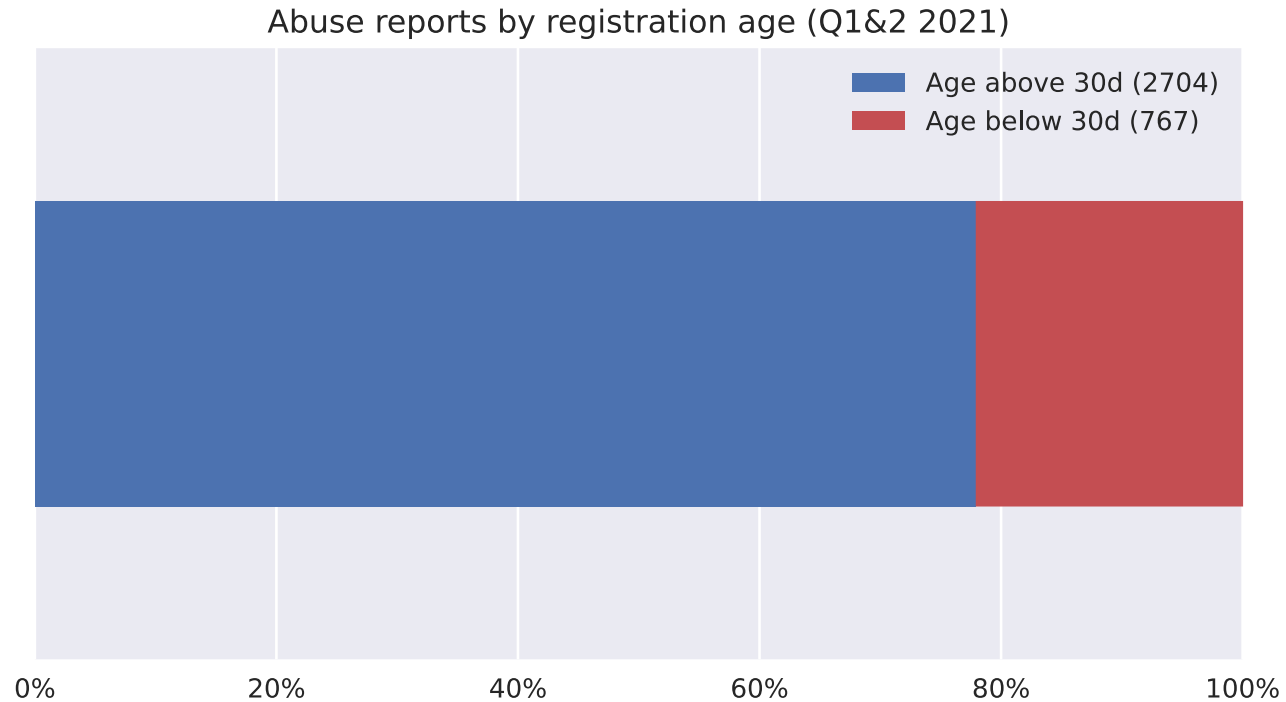
A case study for responsible ML

Thymen Wabeke | ONE Conference

18 October 2022



Abuse regularly involves recent registrations



```
UTC 2022-09-13 20:11:44 (Using a proxy in nl)
Log in bij Mijn ING - ING Bankieren
https://ings[REDACTED].nl/bedankt/index.html
```



```
UTC 2022-09-21 07:04:32 (Using a proxy in us)
Controleer het gebruik van uw account
https://kvk.nl.bed[REDACTED].nl/index.php
```

< KVK.nl



Valideer uw gegevens in h

- Wij vragen u om de actualiteit van uw (contact-)gegevens te cont

Bedrijfsnaam

KvK nummer

Finding the needle in the haystack

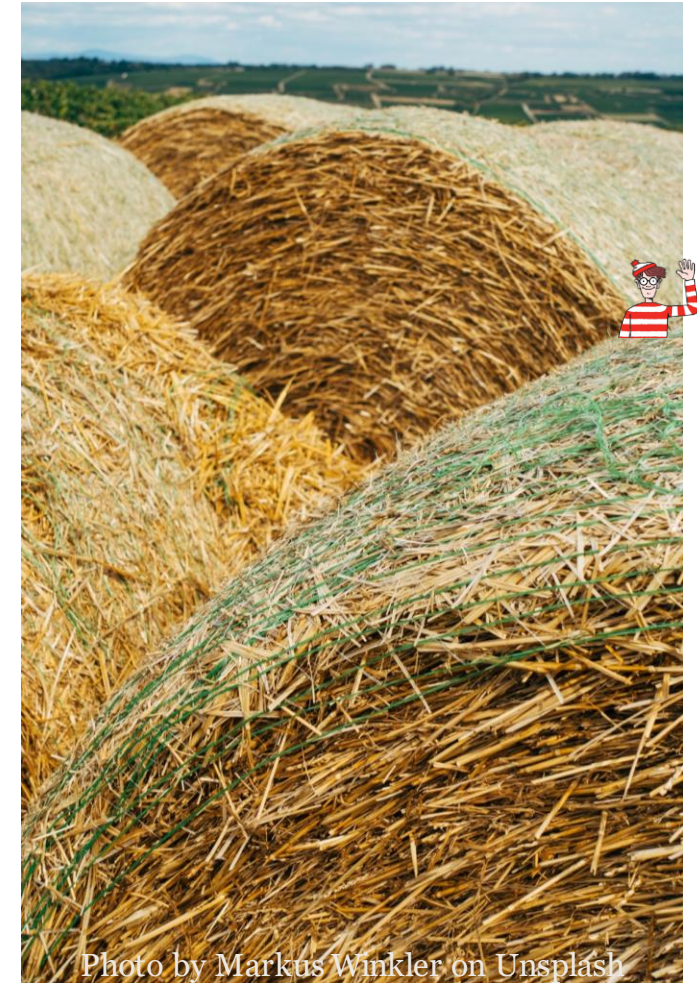
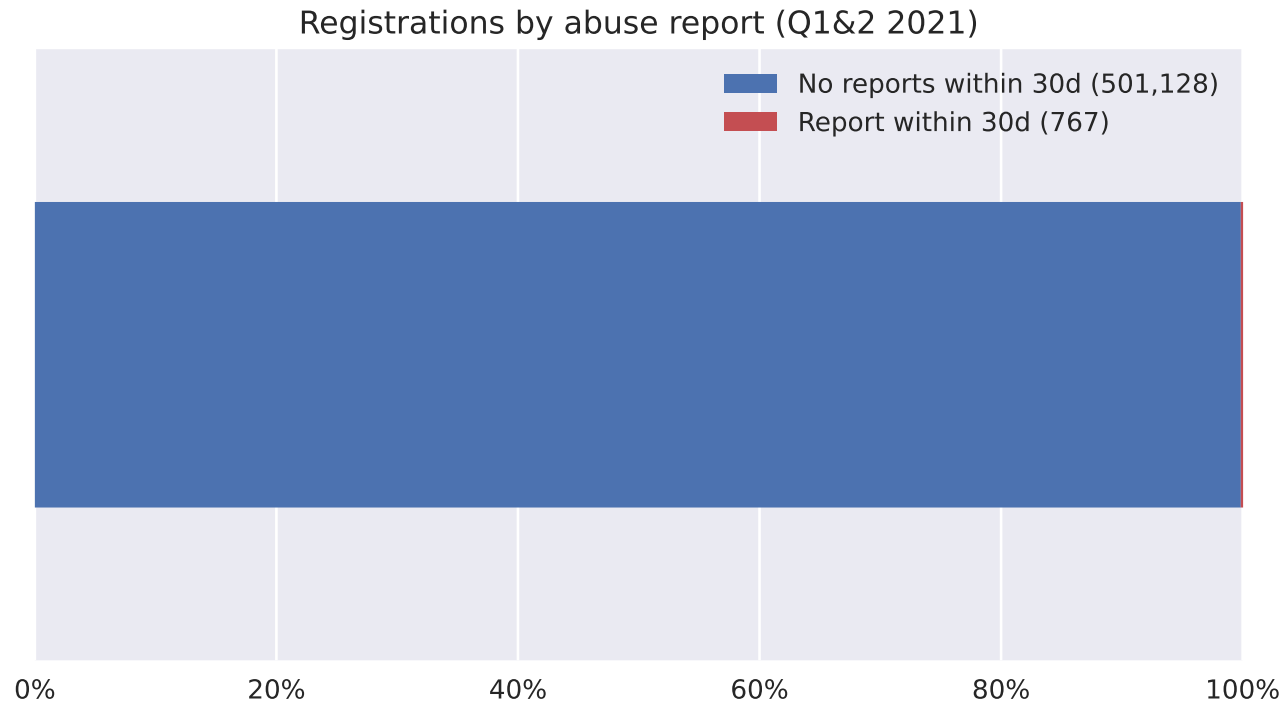
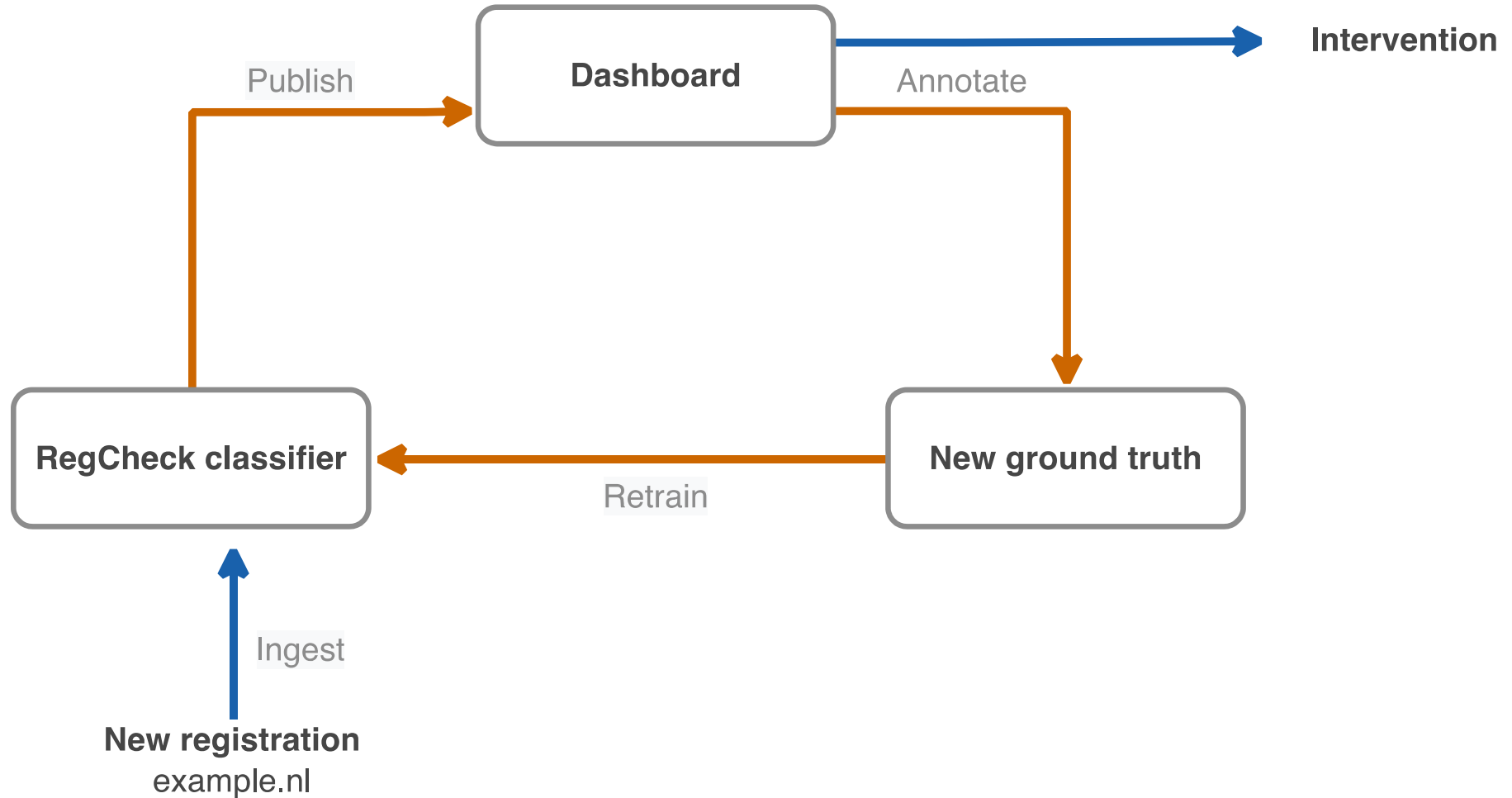


Photo by Markus Winkler on Unsplash

Goal: Identify registrations analysts should review



Results so far

- 3 algorithms
- 3 develop & evaluate iterations
- Various training paradigms
- Current prototype is used daily

Today:

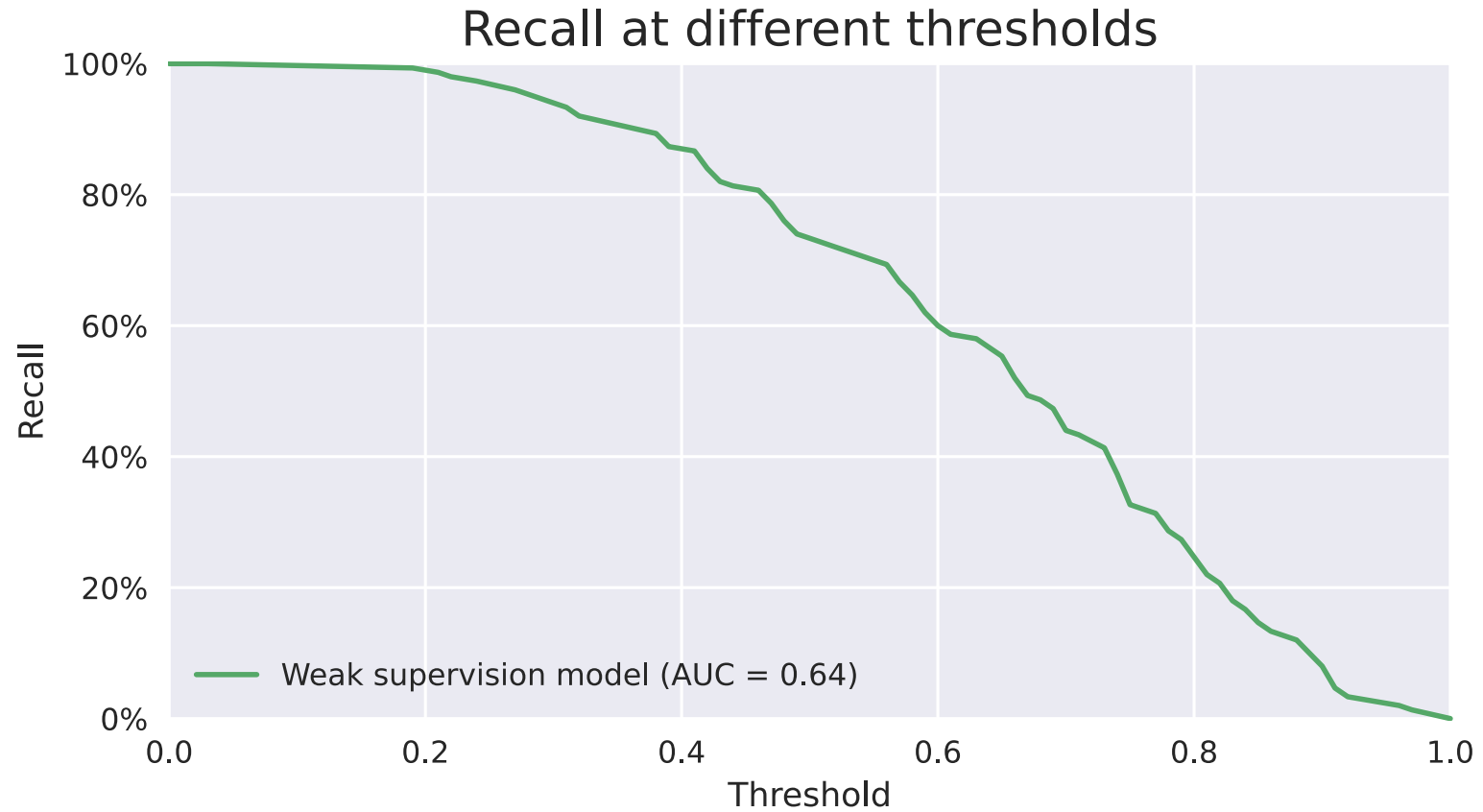
- 4 lessons learned

The screenshot displays a WHOIS lookup for the domain **example.nl**. At the top, there are navigation tabs for WHOIS, DRS Historie, Website, and KASM. The main content is organized into several sections:

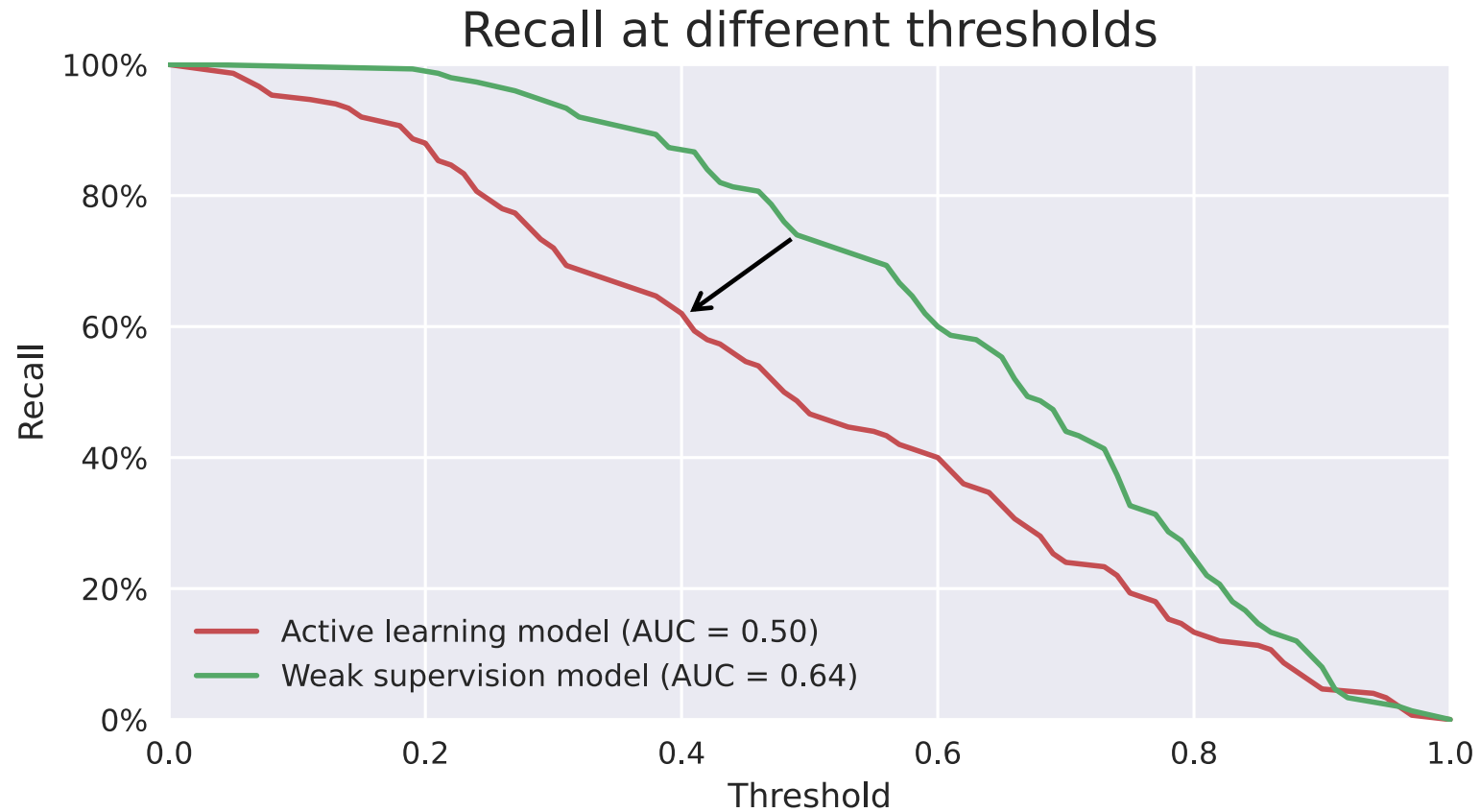
- Zekerheid score:** 0.59
- Houder:** Admin-C
- Name:** Haantje de Voorste
- Address:** Meander 501, 6825 MD, Arnhem, NL
- Email:** sidnlabs@sidn.nl
- Registrar:** The Example Registrar N.V.
- Reseller:** Great Reseller B.V.
- Registration date:** 2022-10-03 06:36:40
- Name servers:** ns1.sidnlabs.nl, ns2.sidnlabs.nl

At the bottom, there is a **Comment** section with a text area containing "Houder ingeschreven bij KvK". To the left of the comment are two buttons: "Reset annotation" and "Previous". To the right, under the heading **Label**, there are two options: "High-risk registration" (checked) and "Registration invalid" (unchecked).

Effect of feedback loop (1/2)



Effect of feedback loop (2/2)



Lesson 1: Speak the same language

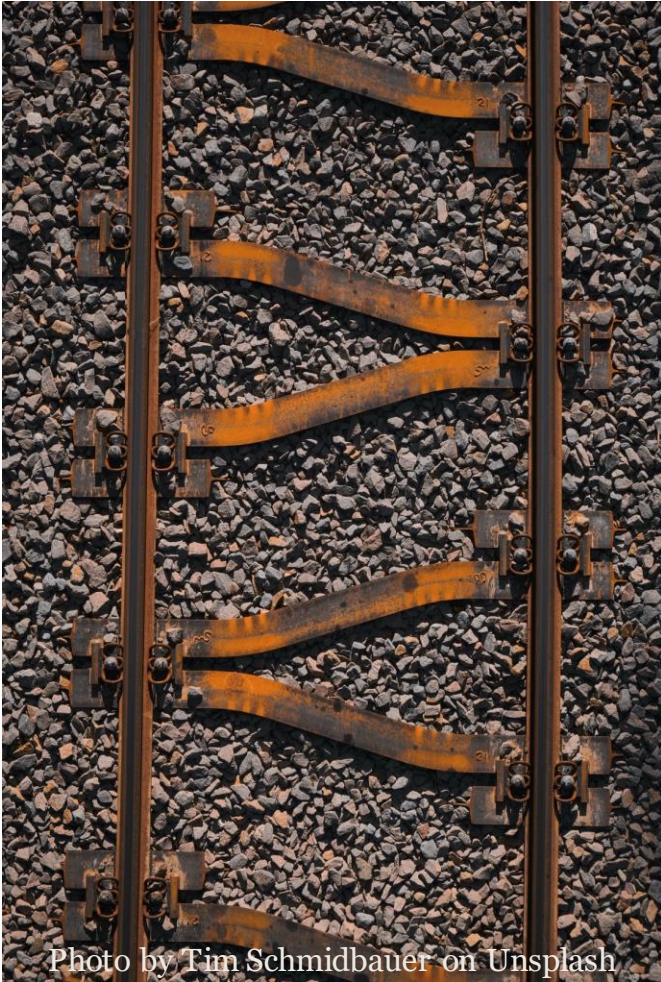
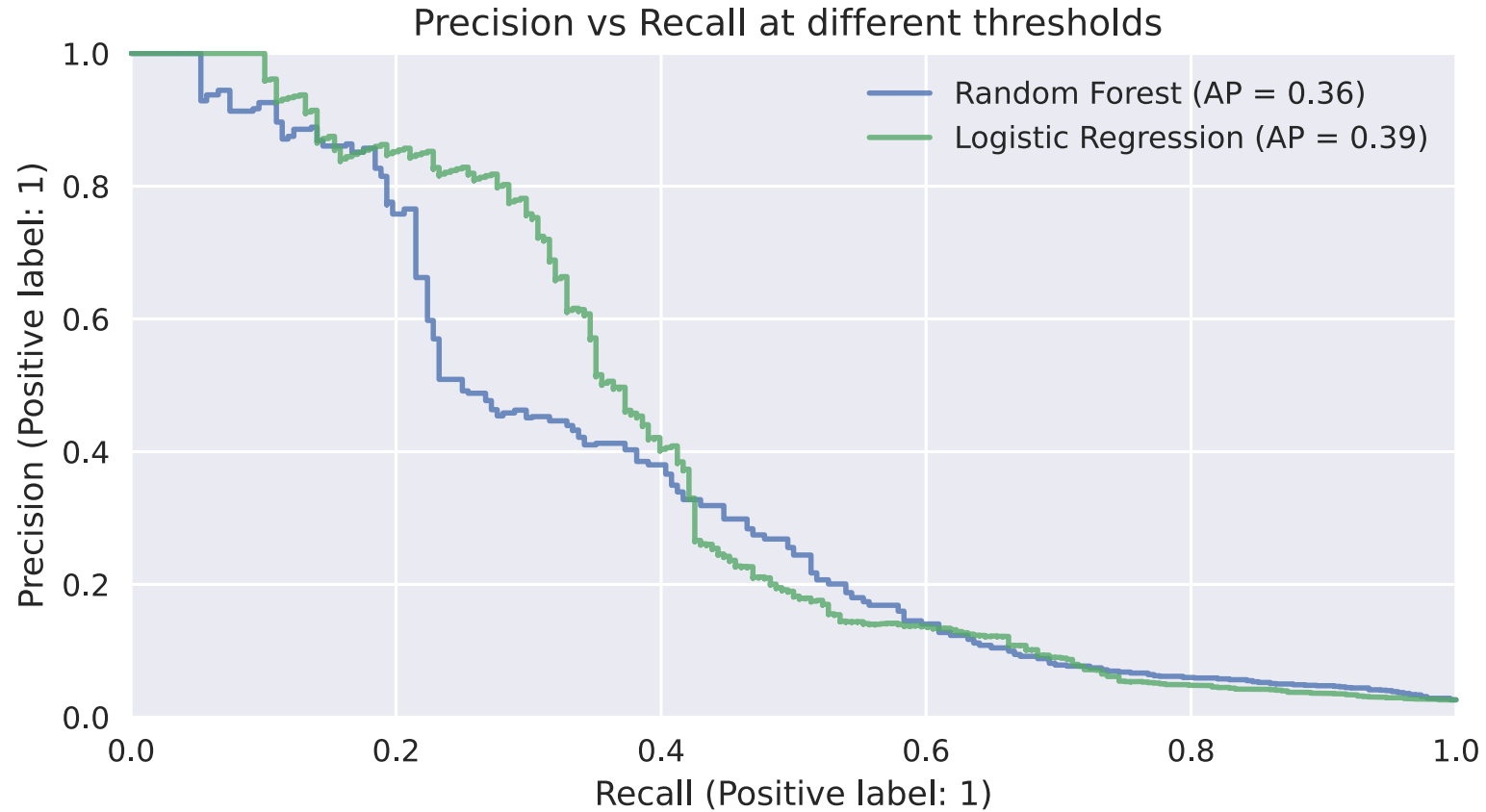


Photo by Tim Schmidbauer on Unsplash

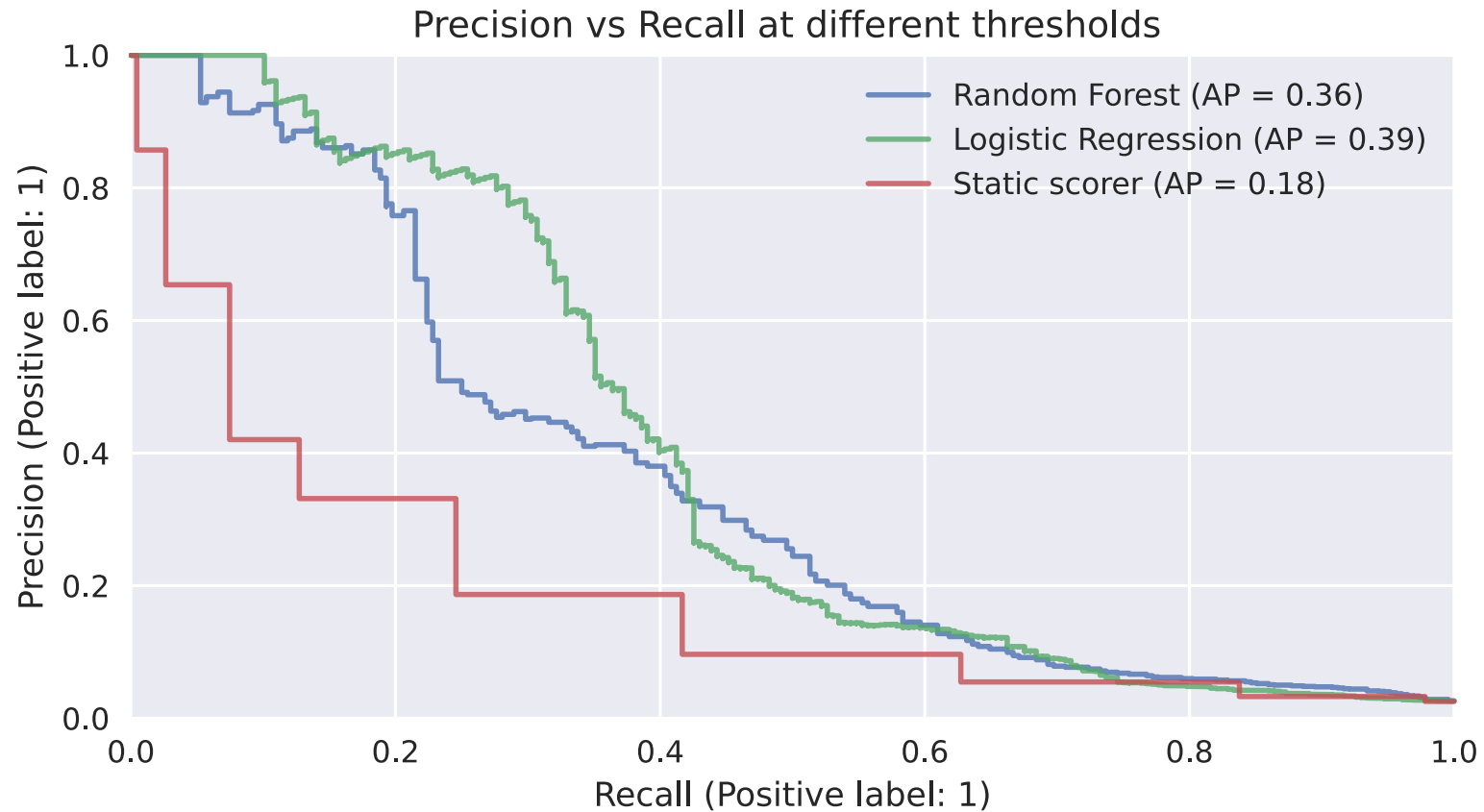
- Problem formulation^[1]
 - Identify high-risk domain names?
 - Identify inaccurate registration data?
- Well-defined outcomes
 - What does “high-risk” mean?
 - Does everyone make the same judgement?

[1] Problem Formulation and Fairness - <https://arxiv.org/abs/1901.02547>

Determine ML performance (1/2)



Determine ML performance (2/2)



Lesson 2: Benchmark ML against non-ML baseline



Photo by Hasan Almasi on Unsplash

- Complex ML algorithms do not necessarily outperform simple ones and may have downsides (e.g., lack of explainability, higher costs) ^[2]
- Adding a non-ML baseline makes pros and cons of using ML explicit

[2] Measuring the predictability of life outcomes with a scientific mass collaboration - <https://www.pnas.org/doi/10.1073/pnas.1915006117>

Evaluation using common ML metrics (1/2)

- Recall = 0.95
(sensitivity)
- Precision = 0.86
- Specificity = 0.85

		Predicted		
		Positive	Negative	Σ
Actual	Positive	95	5	100
	Negative	15	85	100
	Σ	110	90	

Evaluation using common ML metrics (2/2)

- Recall = 0.95
(sensitivity)
- Precision = 0.86
- Specificity = 0.85

		Predicted		
		Positive	Negative	Σ
Actual	Positive	95	5	100
	Negative	15	85	100
	Σ	110	90	

- Expected false negatives = 1
- Expected false positives = 418

	Percentage	Count
Not reported	99.95%	2,786
Reported	0.5%	14
Σ	100%	2,800

Lesson 3: Learn from other disciplines

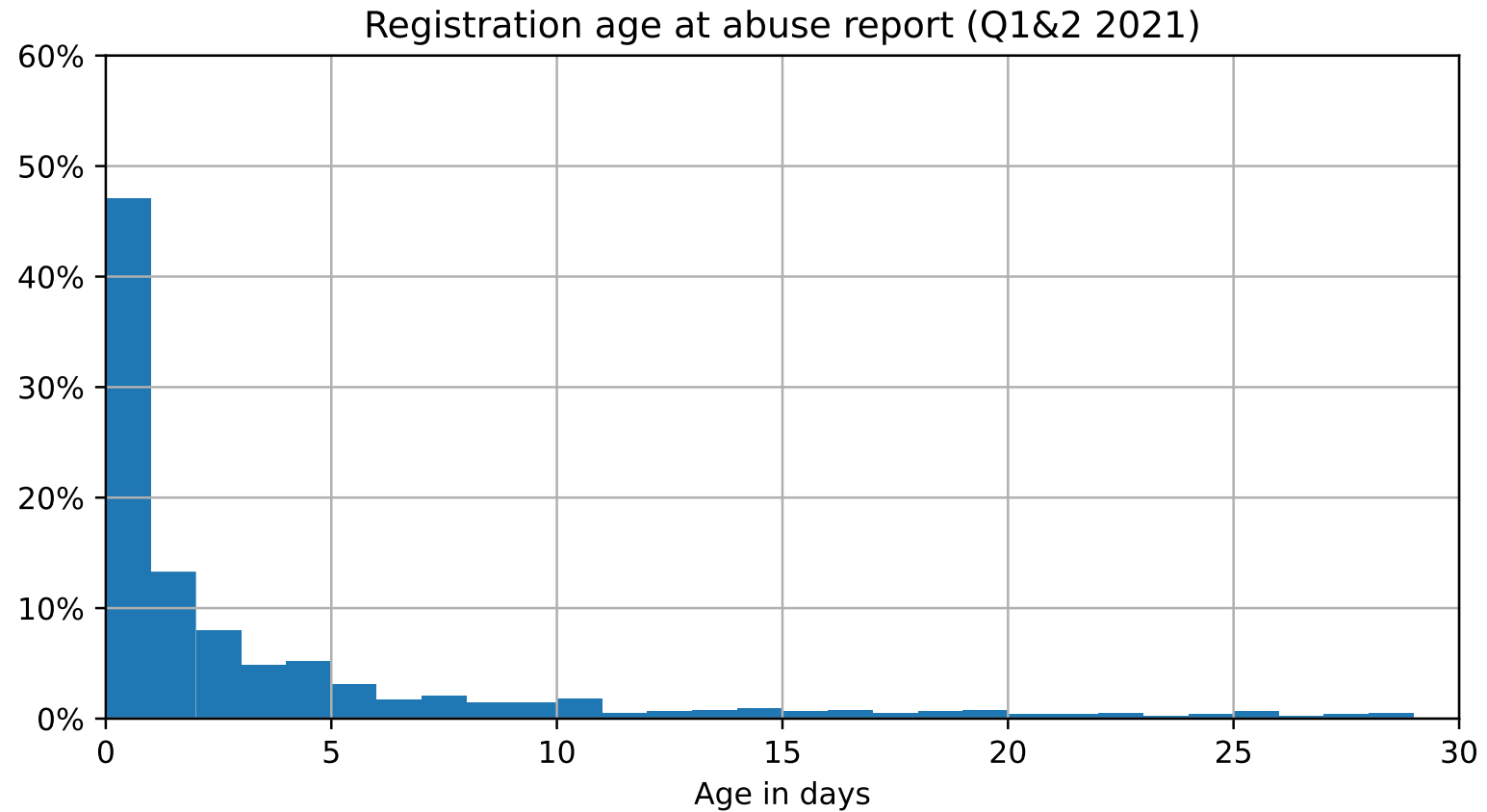


Photo by Tim van Cleef on Unsplash

- Common ML metrics do not tell the whole story, because performance depends on abuse prevalence
- Adopt metrics from medical research: [3]
 - Positive and Negative Predictive Value (PPV, PNV)
 - Pre- and posttest probability

[3] Understanding and using sensitivity, specificity and predictive values - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2636062/>

Accurate vs rapid intervention



Lesson 4: Discuss technical choices and their impact



Photo by Clay Banks on Unsplash

- Technical choices often influence the operation and policy (e.g., threshold, features to include)
- Responsible ML is only possible if decisions are made explicit and discussed from different points of view

4 lessons learned



Speak the same language



Benchmark against non-ML baseline



Learn from other disciplines



Discuss technical choices and impact

Future work

- Improve “operational prototype”
- Joint evaluation and development with DNS Belgium (.be)
- Support peer registries by publishing our methodology

4 lessons learned for responsible machine learning



Speak the same language



Benchmark against non-ML baseline



Learn from other disciplines



Discuss technical choices and impact

Follow us

 SIDN.nl

 @SIDN

 SIDN

Q&A

www.sidnlabs.nl | stats.sidnlabs.nl

Thymen Wabeke
Research engineer
thymen.wabeke@sidn.nl

