In search of solutions against indirect discrimination through AI applications

Digitax Conference on computational taxation: in search for fairness and transparency in tax technology, 23 September 2021

David Hadwick, LL.M.

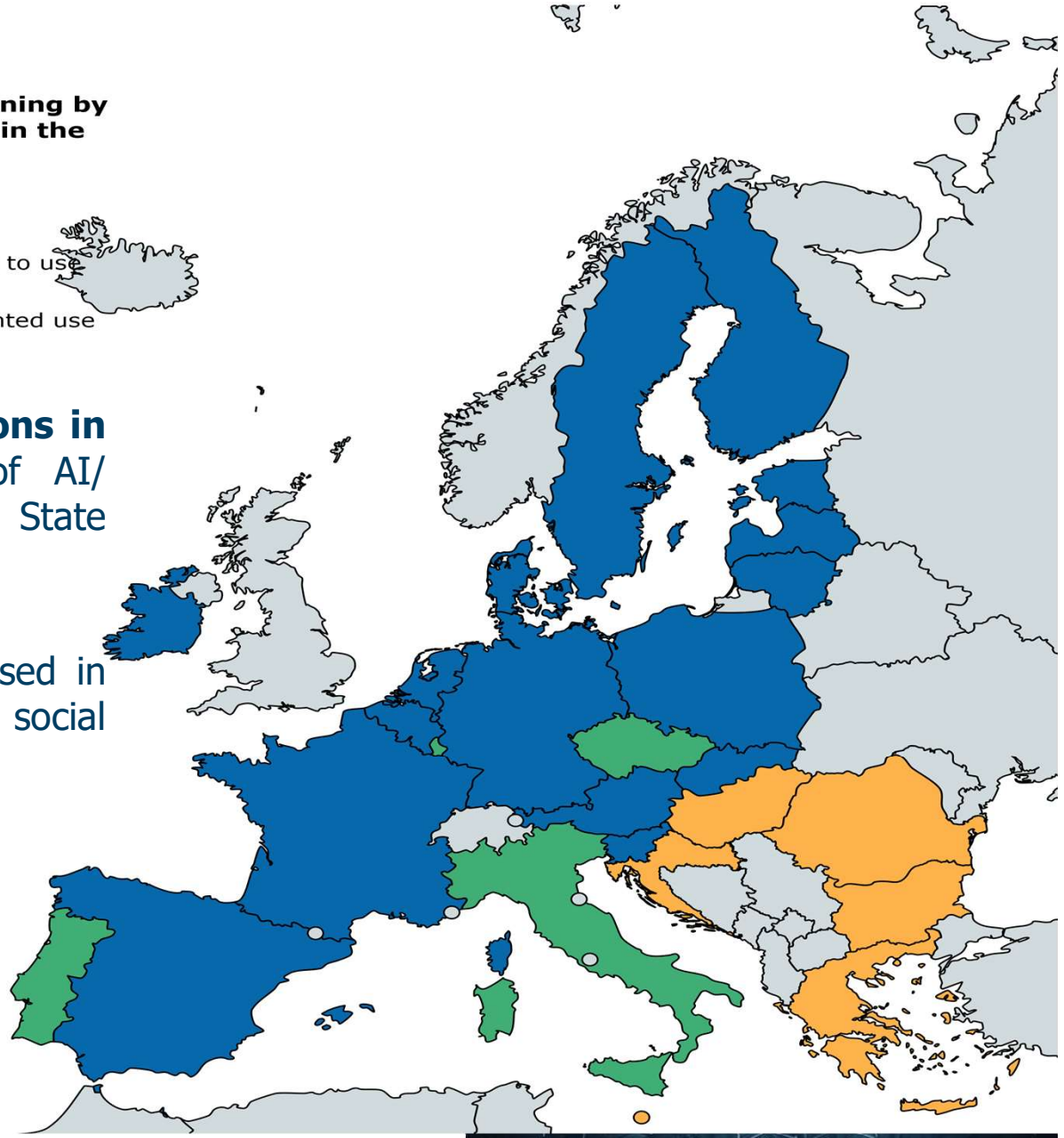
david.hadwick@uantwerpen.be

www.ata-digitax.com

Antwerp Tax Academy
University of Antwerp

Use of machine-learning by tax administrations in the EU

- Tax admin. using machine-learning
- Tax admin. planning to use ML in near future
- No publicly documented use of machine-learning



A majority of tax administrations in the EU (16/27) make use of AI/ machine-learning to perform some State fiscal prerogatives.

Machine-learning is in particular used in the area of VAT, Customs, and social security fraud.

Antwerp Tax Academy
University of Antwerp

DigiTax
Centre of Excellence

EU Law Enforcement

Central point of information, research and discussion

Home Archive Bookshelves Network Student posts About

The Dutch benefits scandal: a cautionary tale for algorithmic enforcement

Björn ten Seldam & Alex Brenninkmeijer

On January 15, the Dutch government was forced to resign amidst a scandal around its child-care benefits scheme. Systems that were meant to detect misuse of the benefits scheme, mistakenly labelled over 20,000 parents as fraudsters. More crucially, a disproportionate amount of those labelled as

SEARCH THIS BLO

Search ...

E-mail

Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

Overall, Northpointe's assessment tool correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend. It makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes. (Source: ProPublica analysis of data from Broward County, Fla.)

Two Petty Theft Arrests

Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.

Antwerp Tax Academy
University of Antwerp

DigiTax
Centre of Excellence

Questions

1. What constitutes prohibited discrimination? What is (in)direct discrimination?
2. How can indirect discrimination arise from the use of AI in taxation?

3. **How can indirect discrimination be avoided?**

1. What constitutes prohibited discrimination?

What is discrimination? (E.g. Article 14 ECHR/Art. 21 EU Charter/Art. 10 Belgian Constitution)

"*A difference of treatment of persons in analogous or relevantly comparable situations, based on an identifiable characteristic or status* (nationality, race, gender, age, religion, etc.)."

Discrimination test – Criteria:

1. **Difference of treatment of persons in … a Comparable/Analogous situation?**
   - Comparator group shows unequal treatment based on protected characteristic
   - Comparability is context-specific, e.g. resident/non-resident

2. **Without reasonable justification? (legitimate aim + proportionality)**

1. Direct v. Indirect discrimination

Direct discrimination:

The difference of treatment is **explicit.**

The measure has **discriminatory intent.**

Is overt, hence easier to detect and to avoid in democratic societies.

E.g.: rule which prohibits ownership for certain nationality/ethnicity.

Indirect discrimination:

general measure, with **neutral terms**, but which has **disproportionately prejudicial effects** on a particular group.

Does not require discriminatory intent.

Is harder to detect and often requires a thorough **ex-post assessment** of the effects of a policy.

E.g.: 'red-lining'.

2. How does indirect discrimination arise in tax?

Indirect discrimination may arise from a **neutral rule** or from a ***de facto* situation.**

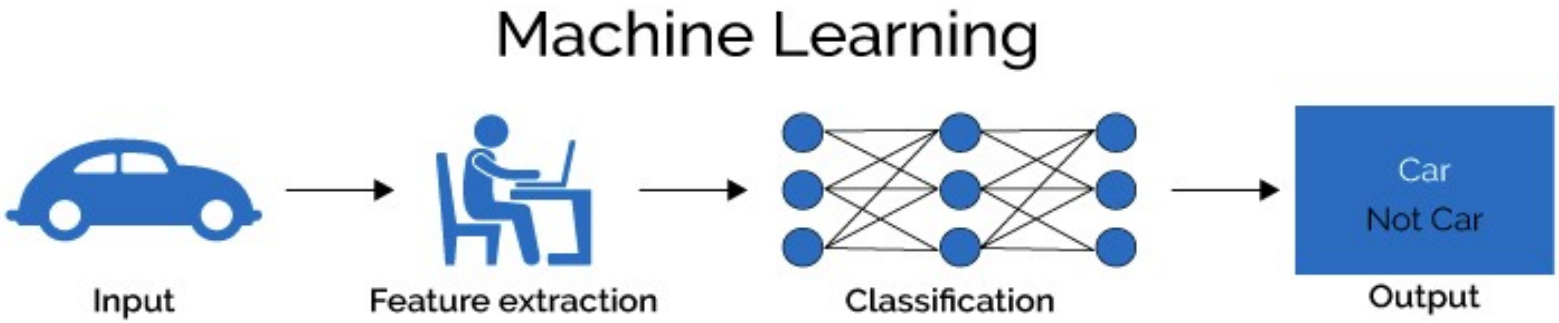
Example: the tax administration decides to audit taxpayers with **high amounts of physical cash**.

Although it is a sound policy, it may have a disproportionately prejudicial effect on **foreign nationals**, because companies with high amounts of cash (e.g. night-shops, catering industry) are in higher proportion foreign-owned businesses.

How can we isolate fraudulent businesses without prejudicial effect on foreign nationals?

How can we avoid adverse outcomes?

2. How does indirect discrimination arise in ML?



Features are extracted from the data, hence ML models should in principle be less biased. (Kahneman et al., *Noise* (2021))

However, if the data is biased, incomplete, erroneous or if the weights attached to certain inputs are incorrect, ML models will recreate or exacerbate existing patterns of discrimination.

2. How can indirect discrimination arise in machine-learning?

Collection/sample bias

Proxies for protected attributes

Under-/overrepresentation of specific groups

Data contaminated with prejudiced cases

Biased training data

Biased target variables and class labels

Biased feature selection

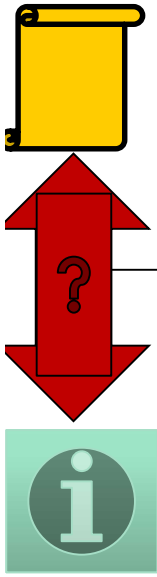
Overfitted models

Systematic/Feedback loops

3. How can indirect discrimination be avoided?

Non-discrimination/data protection norms do not prescribe specific legal or technical safeguards.

Data collection	Training	End-use
Transparency, data accuracy and fairness of data collection.	Transparency of data processing methods.	Non-discrimination, explainability, data subject rights, rights of the defense.
Mechanisms to ensure that taxpayer data is accurate, fair and not tainted with prejudiced cases.	Verifiable insights for taxpayers to understand how their data is processed.	Insights for taxpayers to understand how the machine-learning model arrived at a decision.



Antwerp Tax Academy
University of Antwerp

DigiTax
Centre of Excellence

Conclusion

- The question remains open.
- **Inter-disciplinary dialogue is key!**