

ALGORITHMIC FAIRNESS

University of Groningen
Minor Data Wise 2020/2021
Collaborative Data Project



Goal of the project

This project aimed towards educating people who are unfamiliar with algorithmic fairness or to refresh people's knowledge on this topic. We contributed to the research on algorithmic fairness by modifying a checklist provided by CBS and creating a more extensive checklist focused on fairness. This aids DataFryslân and other future researchers in working and researching within the field of algorithmic fairness as DataFryslân is part of a network which makes the distribution of our work possible.



Deliverables for DataFryslân

- A literature review on the topic of fair algorithms.
- A list of sensitive attributes.
- A checklist which aids data scientists and policymakers in judging the fairness of their algorithm.



Fairness

A fair algorithm prevents unwanted discrimination from happening. Unwanted discrimination can be based on sensitive attributes. The word 'unwanted' is important because sometimes it is 'wanted' to differentiate ('discriminate') based on the sensitive attributes. When doing biological research, for example, gender can be a very useful attribute that is used without doing any harm. Discrimination is 'unwanted' when it causes harm to people.



Sensitive Attributes

Variables which are deemed unfair to base differentiation on. Be aware of proxy variables.

- Gender
- Race
- Age
- Ethnicity
- Relationship status
- Trade-union membership
- Health-related data
- Political opinions
- Genetic/Biometric data
- Religious/Philosophical beliefs
- Sex life/Sexual orientation



Technical Checklist

We provided a technical checklist to aid data scientists on developing fair algorithms. We advise that policymakers should also be included in the process in order to ensure everyone is on the same page.

Main topics

- Exploration
 - Context
 - Laws and regulations
 - Data inventory
- Development
 - Exploratory data analysis
 - Model selection
 - Model performance
- Implementation
 - Apply model
 - Reflection
- Use

**Total of
67
questions**

Example questions

- Have you identified possible risks and benefits of creating/using this algorithm?
- Does the data contain variables that can lead to unwanted discrimination?
- Is there automatic decision-making, whether or not employing profiling?
- Does the model explain the data well enough so that the results can be used for the intended purpose?



Policy Checklist

We provided a policy checklist to aid policymakers in developing fair algorithms. We advise that data scientists should also be included in the process in order to ensure everyone is on the same page.

Main topics

- Context
- Legislation and regulations
 - General Administrative Law Act
 - Public Access to Government Information Act
 - GDPR
- Data inventory
- Model selection and performance
- Implementation
- Use

**Total of
34
questions**

= 101 questions

Example questions

- Has privacy been taken into account, for example by carrying out a Data Protection Impact Assessment (DPIA) to identify privacy risks?
- Have appropriate technical and organizational measures been taken to be able to demonstrate that the processing is carried out in accordance with the GDPR?
- Has the algorithm been audited by internal or external auditors?