

Case follow-up: Putting fairness in context

Jens Ulrik Hansen

Roskilde University

2019-08-29

Case: “Empty Eyes”

- **Exercise 1:**
 - By coincidence, a doctor have discovered that an existing type of medicine, called A, can have a preventive effect on empty eyes. However, this existing medicine is very expensive and short in demand
 - Build a fair model that can decide whether a person should be offered the preventive medicine.
- **How did it go and what considerations did you make?**

Case: “Empty Eyes”

- **Exercise 2:**
 - The government have decided to issue a travel ban on certain people to stop the disease from spreading further
 - Build a fair model that can decide whether a person should be issued a travel ban.
- **How did it go and what considerations did you make?**

Case: “Empty Eyes”

- **Exercise 3:**
 - 10 doctors with equipment to treat Empty Eyes are ready to be deployed around the country.
 - In a fair way decide where to distribute the 10 doctors
- **How did it go and what considerations did you make?**

Case: “Empty Eyes”

- **Exercise 4:**

- The Danish prime minister calls you an early morning in the middle of the department breakfast.
- She tells you delighted that she haven gotten some drug to treat Empty Eyes from her good friend Donald.
- However, there is only enough of the drug to treat 10% of the population.
- The prime minister wants your help to decide which 10% of the population to administer the drug in a way the serves the country best
- Develop an algorithm to select the 10% to which to administer the drug.

- **How did it go and what considerations did you make?**

Case: “Empty Eyes”

- **Exercise 5:**

- A Facebook employee has found a funny cat video that seemingly can cure Empty Eyes.
- Facebook offers to show the video in people newsfeeds, but do not want to spam everybody
- Thus, build an algorithm that Facebook can use to target potential Empty Eyes infected people

- **How did it go and what considerations did you make?**

Case: “Empty Eyes”

- **Exercise 6:**

- Facebook has found out that people with Empty Eyes are bad at clicking adds (as they can't see them).
- Facebook has heard about your ML skills and want to hire you to develop algorithm that can better target users that are likely to click an add.
- However, Facebook is also concerned with their image, so you need to convince them that the algorithm is fair.
- Can you develop such a fair algorithm for Facebook?

- **How did it go and what considerations did you make?**

Exercise 1 versus exercise 5

- Does it matter to fairness whether we are building an algorithm for offering sparse resource (preventive medicine in exercise 1) or an abundant resource (cat video in exercise 5)?
 - **False positives and false negatives**
 - For a sparse treatment/preventive medicine, every false positive is a lost opportunity to save a potential sick person. Thus we want to limit false positives
 - For an abundant treatment, false positives does not matter much (if there are no serious side effects!). As there is no limit to how many we can treat, we might as well try to limit false negatives
 - Is it fair to a person to give them a drug if they are not sick, if the side effects are negligible?
 - Is it fair from a societal point of view?

Exercise 5 versus exercise 2

- Does it matter to fairness whether we are building an algorithm for offering treatment (something positive) or an algorithm for giving people a travel ban (something negative)?
 - **False positives and false negatives**
 - For an abundant treatment, false positives does not matter much (if there are no serious side effects!). As there is no limit to how many we can treat, we might as well try to limit false negatives
 - In the later case, false positives seems to be unfair (to the person that gets a travel ban without carrying the disease)!
 - In the later case, false negatives, might be considered unfair to people that will potentially get infected by a spread of the disease (if the disease is highly contagious and dangerous)

Fairness in context

- Different contexts (different intended use of a ML algorithm) matters to the importance of false positives and false negatives, and fairness in general

Fairness from a point of view

- Different stakeholders might have different views on what matters and what is fair
 - From the individual point of view, people might want to have a fair chance of being offered preventive treatment and they might want not to be given an unfair travel ban
 - From a societal point of view, we might want to prevent the disease from spreading and thus it seems fair to put a travel ban on all individuals we expect to have the disease
- Not the intended use of a ML algorithm matters, but also what stakeholder perspective we take

Fairness with respect to what groups

- Is it fair to include gender as a predictor variable, if we are building an algorithm for diagnosing a serious disease?
- Is it fair to include gender as a predictor variable, if we are building an algorithm to predict who clicks an add?

Fairness is not all that matters

- Note several of the previous considerations was not only about fairness
- Other things matters as well from a legal or ethical point of view
 - Preventing a disease from spreading, high societal well-being
 - Being fair to groups (geographically as in exercise 3) or treating as many people as possible
 - How do we select whom to treat, if we cannot treat everyone, we predict to be sick? (exercise 4)

Conclusion

- Fairness is super hard!
 - Technically, there are many plausible and inconsistent definitions
 - In addition, the context and relevant stakeholders matter for what can be considered fair
 - How are we to be fair then?...