Machine Learning CSE2510 – Lecture 5.2: Bias in Machine Learning

Filippo Santoni De Sio

Guest lecturer

TBM

F.SantoniDeSio@tudelft.nl

TU Delft

Welcome to week 5 - lecture 2

- Questions?
- Recap previous lecture
- How to read a scientific paper?
- Read article on bias, while answering 2 questions

<break>

- Discussion on article

Questions?

Recap of the previous lecture

Implicit bias

- A preference or inclination for or against something
- Based on learned coincidences, which unknowingly affect everyday perceptions, judgment, memory, and behavior
- Subconscious thought
- We all have it
- Might result in discrimination

Multiple sources of bias in ML

- Training data
- Lack of diversity in ML developers
- Implicit human biases in our culture
- Evil programmers

Debiasing ML → Fairness in ML

- Fairness is a multi-faceted concept

To improve fairness:

- Debias training data
- Use unbiased features
- Build smart algorithm

But: is difficult to ensure on both the group and individual level simultaneously

Main point

- Building fairness and non-discriminatory behaviour into AI models is not only a matter of technological advantage but of social responsibility

Today's learning objectives

After practicing with the concepts of today's lecture you are able to:

- Critically read a scientific paper

- Argue about and pinpoint bias in a system/organisation taking into account technical, societal, legal, and/or educational point of views

How to read a scientific paper

- 3 passes (i.e., 3 reads):
  - Each pass accomplishes specific goals and builds upon the previous pass

- Pass 1: General idea about the paper
- Pass 2: Grasp the paper's content
- Pass 3: Understand the paper in depth/details

[From: S. Keshav (2007). How to read an article. ACM SIGCOMM Computer Communication Review archive, 37(3), pp. 83-84]

Pass 1: General idea [5-10 mins]

1. Read the title, abstract, and introduction
2. Read the section and sub-section headings
3. Read the conclusions
4. Skim through the references

After Pass 1

You should be able to answer:

1. What type of paper is this?
2. Which other papers is it related to? Which theoretical bases were used to analyze the problem?
3. Do the assumptions appear to be valid?
4. What are the paper's main contributions?
5. Is the paper well written?

Pass 2: Content [1 hour]

1. Read with greater care, but ignore details, e.g., formulas and proofs
2. Look carefully at figures, diagrams, illustrations
3. Mark relevant unread paragraphs for pass 3

After Pass 2

You should be able to summarize the main thrust of the paper, with supporting evidence, to someone else

Pass 3: Details

1. Attempt to virtually re-implement the paper: that is, making the same assumptions as the authors, re-create the work
2. Identify and challenge every assumption in every statement
3. Think about how you yourself would present a particular idea

After Pass 3:

You should be able to

1. Reconstruct the entire structure of the paper from memory
2. Identify its strong and weak points
3. Pinpoint implicit assumptions
4. Identify potential issues with experimental or analytical techniques

Read the article

- Natasha Singer (July 26, 2018): *Amazon's Facial Recognition Wrongly Identifies 28 Lawmakers, A.C.L.U. Says*

https://www.nytimes.com/2018/07/26/technology/amazon-aclu-facial-recognition-congress.html

Background info – the article by A.C.L.U:

https://www.aclu.org/blog/privacy-technology/surveillance-technologies/amazons-face-recognition-falsely-matched-28

Questions

1. Where in the organisation does bias occur?

2. How could you prevent, solve, and/or migitate the bias? Think about this from a technical and/or a societal and/or a legal and/or educational point of view.

Write down the answers individually

Time for a break

Discussion time

15 mins

1. Organise yourselves into groups of 4/5 people
2. Compare and discuss your answers to the questions in your group
3. General discussion

Compare and discuss the questions

1. Where in the organisation does bias occur?

2. How could you prevent, solve, and/or migitate the bias? Think about this from a technical and/or a societal and/or a legal and/or educational point of view.

General discussion

- Where in the organisation does bias occur?

General discussion

- How could you prevent, solve, and/or migitate the bias?
  - Technical point of view
  - Societal point of view
  - Legal point of view
  - Educational point of view

Additional discussion questions

- What happens when the similarity score is lowered?

- What is the expected effect of changing the similarity score on bias?

- What can you say about Amazon's response that:

"A.C.L.U. had used the company's face-matching technology, called Amazon Rekognition, differently during its test than the company recommended for law enforcement customers"?

One solution to the bias in the system that was proposed was:

“to hire “more lawyers, engineers and data scientists of color to assist in properly calibrating this technology to account for racial bias that can lead to inaccuracies with potentially devastating outcomes.” ”

- Why would that help?