Ethical Issues in using Machine Learning and AI

Mahdi Roozbahani
Georgia Tech

These slides are from **Nakul Gopalan** at Georgia Tech

A Few Notes (From Nakul’s notes)

- I am trying this for the first time the topic itself is nascent in its development
- Topics might include ideas of discrimination, bias, ethnicity
- Do not assume malice, instead assume people are making arguments in good faith!!

Outline

- Success of Machine learning/ AI
- Case studies in negative effects of ML:
  - Job Losses
  - Privacy
  - Bias
  - Manipulation with ML
- Ethical considerations in other sciences
- Solutions?
  - Explainability
  - Bias reduction
  - Novel class of algorithms

Outline

- Success of Machine learning/ AI
- Case studies in negative effects of ML:
  - Job Losses
  - Privacy
  - Bias
  - Manipulation with ML
- Ethical considerations in other sciences
- Solutions?
  - Explainability
  - Bias reduction
  - Novel class of algorithms

The Wins

- Natural Language Processing
- Search
- Forecasting: Energy/ Weather/ Finance/ Covid
- Autonomous driving/ Submarines/ Drones
- Learning Controllers - Nest
- Drug discovery/ automated image analysis
- Game Playing
- Traffic prediction

ML applications in social good

- Forecasting severe weather events: US, Australia, Floods in Asia etc.
- Tracking and protecting wildlife: in Tanzania, marine wildlife
- Rural water supply
- Helping blind people sense
- Detecting plant diseases
- Other healthcare based applications
- https://www.mckinsey.com/featured-insights/artificial-intelligence/applying-artificial-intelligence-for-social-good
- https://ai.google/social-good/
- Series of workshops about AI for social good: https://aiforsocialgood.github.io

Outline

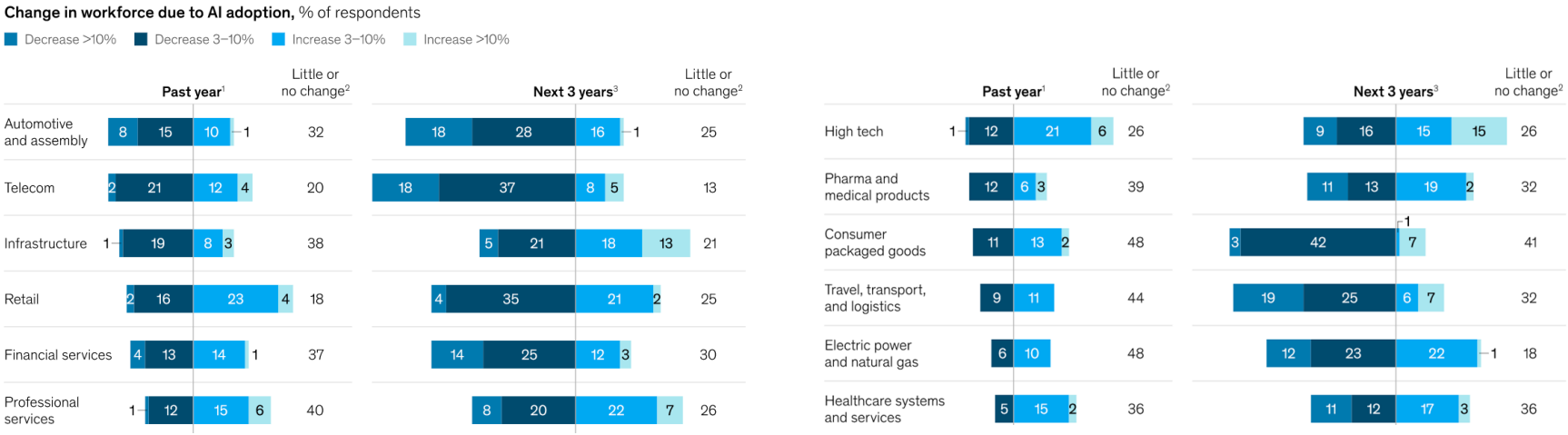
- Success of Machine learning/ AI
- Case studies in negative effects of ML:
  - Job Losses
  - Privacy
  - Bias
  - Manipulation with ML
- Ethical considerations in other sciences
- Solutions?
  - Explainability
  - Bias reduction
  - Novel class of algorithms

Job Losses

- Trucking
- Manufacturing
- Clerical and Secretarial jobs
- Coding???

Jobs at risk of automation in the EU 19 countries

Executives from 1,872 enterprises worldwide report the largest AI-induced workforce contraction in automotive and assembly and telecoms in the last year. Looking forward, the CPG, transport, utilities, retail and financial services are expected to follow.



Job Losses

The loom unemployed mill workers??? (Increasing the output per worker)

Spreadsheets caused job losses amongst accountants???

Will governments step in and help??

What do you think?

Privacy: Cookies

- Easy tracking of people with anonymized IDs
- Selling these fingerprints for advertisements
- Apps collecting data without knowing what it will be used for: Normalized by major ML companies
- General Data Protection Regulation (GDPR):
  - Laws about what can be collected
  - Default settings
  - Sharing

Privacy: Images

- Creating customer IDs from images
- Images on apps like FB
- Tracking customers inside and near showrooms
- Traffic cameras
- Surveillance in oppressive governments – cause of major concern
- Amazon/ UPS/ large mail carriers tracking employee behavior

Privacy: Guiding behavior

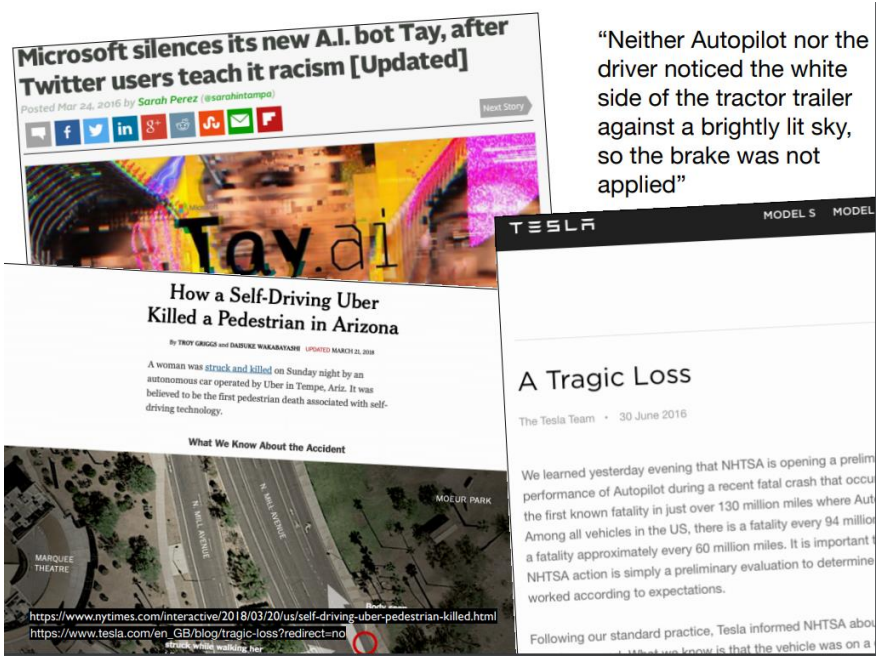
- Advert clicks
- Manipulating user’s feed based on behavior / Wiki link
- OKCupid: We experiment on human beings
- Uber/ Lyft

Privacy

What do we think????

Bias

- We use ML in criminal sentencing
- Biased data in Biased decisions out
- We us ML in deciding lending rates
- Biased data in Biased decisions out
- Bias in large language models



Bias

One of the larger problems in ML

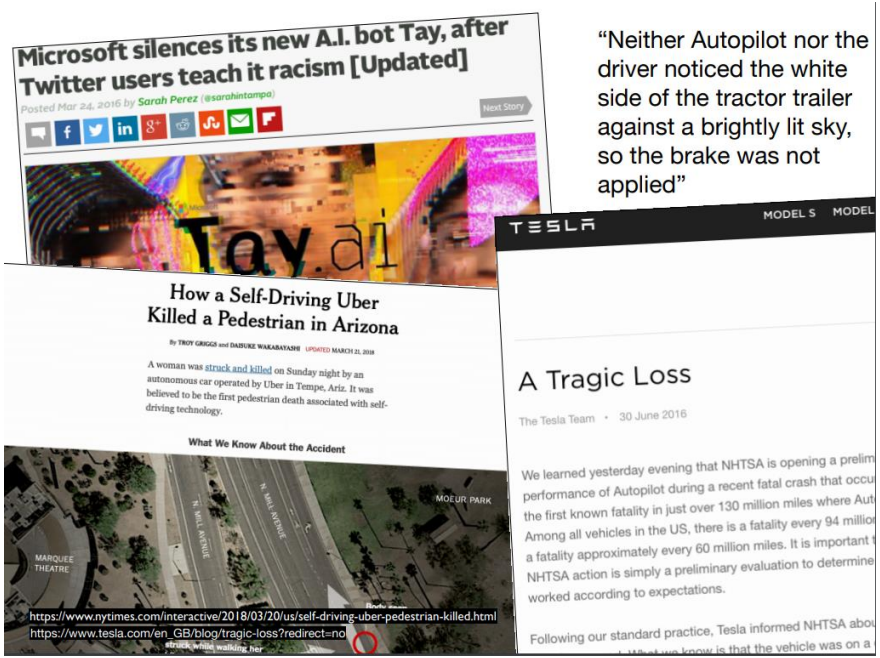
Overfitting and Bias are serious problems with real life consequences

Bias

What do we think?

Manipulation with ML

- Text generation for News
- Picture/ Video Manipulation - Deep fakes



Some deep fake examples:

Manipulation with ML

What do we think?

Ethical risks: A group of researchers have spent years helping to frame the ethical risks of deploying ML in certain sensitive contexts. This year those issues went mainstream.

Examples include policing, the judiciary and the military. A few trailblazing researchers include:

- *Joy Buolamwini, Timnit Gebru*, Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification (2018)
- *Clare Garvie, Alvaro Bedoya, and Jonathan Frankle*. The Perpetual Line-Up: Unregulated Police Face Recognition in America (2016)
- *Adam Harvey*. Megapixels (2017)
- *P Allo, M Taddeo, S Wachter, L Floridi*. The ethics of algorithms: Mapping the debate (2016)
- *Margaret Boden, Joanna Bryson, Alan Winfield* et al. Principles of robotics: regulating robots in the real world (2017)

Outline

- Success of Machine learning/ AI
- Case studies in negative effects of ML:
  - Job Losses
  - Privacy
  - Bias
  - Manipulation with ML
- Ethical considerations in other sciences
- Solutions?
  - Explainability
  - Bias reduction
  - Novel class of algorithms

Medicine

- Ability to do harm

➢ Hippocratic Oath

Genetics

- Eugenics
- Human genetics experiment

➢ Laws, but after a while

Finance

- Fiduciary responsibility
- Insider trading

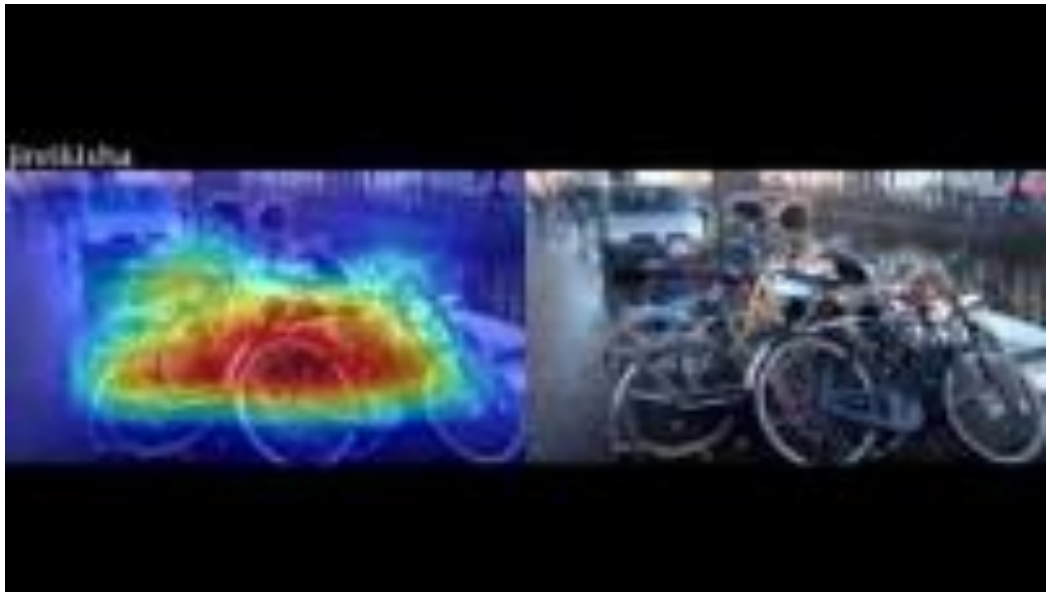
➢ Financial regulation

Outline

- Success of Machine learning/ AI
- Case studies in negative effects of ML:
  - Job Losses
  - Privacy
  - Bias
  - Manipulation with ML
- Ethical considerations in other sciences
- Solutions?
  - Explainability
  - Bias reduction
  - Novel class of algorithms

Explainability

- Explain why?
- Modern machine learning algorithms do not explain their decisions (We have seen it in CNN)
- Approaches:
  - Examples
  - Natural language
  - Counter-example
  - Causal
  - CNN visualization



Bias reduction

- How to reduce bias?
- Better data
- Algorithmic bias reduction

Novel class of algorithms?

- Someone is already working on these?
- Links of interest:
  - Charles Isabell and Michael Littman talk at Neurips 2020
  - Ayanna Howard’s course on AI, Ethics and Society

Summary

- ML/AI are important tools
- Can be used for good
- Can be used to do bad
- Important to consider the end-use of your algorithms
- Important to consider bias and overfitting issues
- Important to understand how your data collection will be used in future