ETHICAL ISSUES IN MACHINE INTELLIGENCE

SPRING 2022

DR. GARRETT DANCIK

KEY ASPECT OF MACHINE INTELLIGENCE

- In machine learning, machines learn from data

- Learning involves updating model parameters in a way that minimizes a cost function (optimization based on a goal)

- Machine learning algorithms may be complex, raising questions about interpretability

- Machine learning algorithms may be not open (code may be proprietary)

MACHINE LEARNING MAY PERPETUATE HUMAN BIASES

- "Silly example": https://arxiv.org/pdf/1602.04938.pdf

- Amazon scraps secret AI recruiting tool that showed bias against women, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G

- Racial bias in recidivism prediction, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

WHO IS ACCOUNTABLE FOR MACHINE-BASED DECISIONS?

- Whose fault is it when self driving cars get into an accident?

- Who is responsible if a machine misdiagnoses a patient?

THE TROLLEY PROBLEM (CLASSIC VERSION)

There is a runaway train (trolley) barreling down the tracks, upon which 5 people are tied to the tracks and unable to move. The trolley is headed straight for them. Nearby is a lever that will switch the train to a different set of tracks, where one person is tied up. What do you do?

(A) Do nothing: the train kills the five people on the main track

(B) Pull the lever: the train changes tracks and one person is killed.

THE TROLLEY PROBLEM (FAT MAN VERSION)

There is a runaway train (trolley) barreling down the tracks, upon which 5 people are tied up and unable to move. The trolley is headed straight for them. You are on a bridge above the tracks, next to a very large man. What do you do?

(A) Do nothing: the train kills the five people on the main track.

(B) Push the man off the bridge and onto the track: the man will be killed, but the five people will be saved.

THE TROLLEY PROBLEM FOR SELF-DRIVING CARS

- MIT moral machine: http://moralmachine.mit.edu/

- The Social Dilemma of Autonomous Vehicles: https://www.researchgate.net/publication/301293464_The_Social_Dilemma_of_Autonomous_Vehicles

DEEP FAKES, AUDIO AND VIDEO MANIPULATION

- What limits, if any, should be put on audio and visual creation and manipulation?

  - https://www.businessinsider.com/dangerous-deepfake-technology-spreading-cannot-be-stopped-2019-7

DIGITAL ETHICS AND THE FILTER BUBBLE

- Machine Intelligence makes human morals more important

  - https://www.youtube.com/watch?v=hSSmmlridUM

- Digital ethics and the future of humans in a connected world

  - https://www.youtube.com/watch?v=bZn0IfOb61U

- The Filter Bubble

  - https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles