

## Identifying Bias in Predictive Algorithms: Facial recognition

Learning objectives: students will be able

- a. to identify and describe bias in machine learning products
- b. Explain the impact of bias on different groups and users
- c. Create criteria for public deployment of ML products

Scenario: A company has developed an algorithm that can detect the same person in different pictures. It works very well but occasionally fails.

**Task 1** Individually, think of reasons or groups of users for which the system can fail (2 min). Fill in poll everywhere to compare answers (2 min)

**Task 2** In a group, for each group/scenario, identify the impact on the person who has been misrecognized. Select and rank the top three most problematic failures and justify this selection (10 min)

**Task 3** Debate with the other group at your table, which failure scenario is most problematic (10min)

Class discussion of the top concerns (10 min)

Lecture (15 mins)

**Task 4** You are a technical advisor to a police unit. The police chief asks you to choose between different providers of facial recognition technology and a recommendation if they should deploy it. Devise a plan on how to test the systems and criteria for deciding if the system is ready for deployment.

**Task 5** One minute paper “What did I learn today”

Homework (individual): For speech recognition systems (i) identify groups or situations when they can fail (ii) identify the impact of errors on users (iii) propose evaluation criteria to choose between systems