Introduction to Algorithmic Bias

Overview

Using machine learning, computers are able to use large datasets to teach themselves. This is often much easier than trying to teach a computer directly, and it's led to computers being able to recognize complex things! However, when we let computers learn from data, it becomes very important to choose the right data to ensure they learn the right thing. Algorithmic bias can happen when computers are given incomplete or biased sets of data to learn from; this can cause them to arbitrarily prioritize one group or idea over another.

Activity Plan

Part 1: Examples of "Great Art"

Say: *We are going to be teaching a computer to recognize great art! To do that, we're going to put together a database of examples of great art for the computer to learn from. It will use these to identify other great art.*

Lesson Goals

- Explain algorithmic bias
- Identify why algorithmic bias exists and how it happens
- Explain how we can prevent algorithmic bias

Materials Needed

- Access to the internet or other way for students to find examples of great art
- Great art datasets created by experts (link)

Have students put together a set of 20 examples of great art. They can work alone or in small groups, but don't let the groups get too big--2 or 3 students at most.

Part 2: Comparisons

When all the students have put together a group of images, have them share their datasets. Then show the students the datasets that were put together by the following experts:

- Google - these are the top 20 hits for a search for "great art"
- Historylists.org - a website with a mission to bring history closer to people
- Maureen - studied art history, practicing visual artist
- Dan - studied architecture, works as an architect
- Jibran - went to art school, works as a product designer
- Shezza - studies culinary arts and Islamic art

Datasets: https://tinyurl.com/anp-AlgorithmicBias

Part 3: Discussion

- What are the differences between these datasets? How would you describe each one?

- What would have happened if you'd used one of these datasets to train the computer to recognize great art? What would have been included? What would have been left out?

- Which dataset/combination of datasets would you use as your final training data for the computer?

- What do you think we should keep in mind when we're creating datasets for machine learning algorithms?

> *We have pale male datasets being used as something that's universal, when that isn't actually the case.*
>
> *-Joy Buolamwini*

Further Learning

Part 4: Algorithmic Bias in Software

Have students watch Joy Buolamwini's
ideo AI, Ain't I a Woman? and/or some of the following resources on

issues with algorithmic bias:

- Videos:
  - AI, Ain't I A Woman?, Joy Buolamwini
  - Joy Buolamwini: How I'm fighting bias in algorithms
  - AOC & Buolamwini clip from House Hearing on Facial Recognition Technology
  - Coded Bias film
- Articles:
  - Wrongfully Accused by an Algorithm, New York Times
  - Algorithmic Bias: Why Bother?
  - Man is to Doctor as Woman is to Nurse: the Gender Bias of Word Embeddings
  - Companies are on the hook if their hiring algorithms are biased
- Books:
  - Technically Wrong by Sara Wachter-Boettcher

Part 5: Discussion

The American government is currently trying to figure out if facial recognition software should be allowed to be used by the government and police. They are also thinking about how and if they should regulate algorithms. What do you think?

- Should we allow facial recognition software?

- Are there times when it's ok and times when it's not ok?
  - Should we use facial recognition software to try on glasses on the internet?
  - Should we use facial recognition software to identify suspects in a crime?
  - Should we use facial recognition software from the security cameras all over town to try to locate a missing person?

- Is it possible to prevent algorithmic bias?

- What steps can we take to try to make sure the software we create is equitable and positive for everyone?