

Tuesday 7/26 - Bias, what is it?



Important link with more important links:

Announcements:

-

In the News:

-

Agenda:

- What is AI?
- AI Bingo (**EXTRA CREDIT OPPORTUNITY!!**)
- Bias

Resources

- [The Secret Bias Hidden in Mortgage-Approval Algorithms](#)
- [Amazon scraps secret AI recruiting tool that showed bias against women](#)
- [The Economic Exclusivity of Trader Joe's](#)
- [Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms](#)
- [Why algorithms can be racist and sexist](#)

Songs of the Day

- Oogum Boogum Song - Brenton Wood
- Vanille Fraise - L'impératrice
- Might be - Anderson.Paak

- Soull II Stro - Stro Elliot

Do not edit above this line.

Questions or Things to Share?

-
-

INDIVIDUAL REFLECTION

What is an artificial intelligence (AI)? Write your name and thoughts below

AI Bingo

On your bingo card there are boxes with different AI systems. Your task is to go around the room and partner up with someone and identify the data used and the prediction made by the AI system.

The first student to get bingo (5 in a row, column, or diagonal) will get 2 extra credit points

BIAS? I'VE NEVER HEARD OF HER!

Can an AI be biased? How so? Give some examples. Write your name and responses below.

I CAN'T BE BIASED, I DON'T SEE COLOR!

It is well documented that people of color, especially Black and Brown people, are denied mortgage loans far more often than white people. Fannie Mae and Freddie Mac, two government backed mortgage provider services, claim that their algorithm they use to determine eligibility for mortgages is "color-blind" since they do not consider a person's race in the application. Here is the data that Fannie May and Freddie Mac do use:

- credit histories
- the prospective borrowers' assets
- employment status
- Debts
- the size of the loan relative to the value of the property they're hoping to buy

Is Fannie Mae and Freddie Mac's algorithm truly color blind? Do you think the algorithm is biased? How would you design the mortgage approval algorithm to be less biased?

SO WHAT NOW?

How might you mitigate/address the following biases if you were building an ML model?

- Historical Human Bias
- Incomplete or unrepresentative training data

Write your name & reflection below

IS BIAS BAD :(

Is bias always a bad thing? Discuss with your group

FINAL REFLECTION

Can AI systems eliminate bias? How? Why or Why not? Write your name and answers below.

READING/WATCHING FOR TONIGHT

Wednesday 7/27 - Fair!? For Whom?



Announcements:

-

In the News:

Agenda:

- Individual reflection
- COMPAS and Fairness Metrics
- Loan Discrimination

Additional Resources

- [Elements of AI free online course](#)
- [Classification: True vs. False and Positive vs. Negative](#)
- [Google's Machine Learning Education Hub](#)

Songs of the day:

- Who You Foolin - Gunna
- Far Away - Beast Coast, Joey Bada\$\$, Flatbush Zombies
- Tearin' up My Heart - 'N Sync

Do not edit above this line.

INDIVIDUAL REFLECTION

What does fairness mean to you? Write down your name and response below

SHE IS VERY FAIR TO ME!

In the ProPublica COMPAS article you read last night, the main issue was that the COMPAS algorithm designer, Northpointe claimed that their algorithm was fair whereas ProPublica claimed that it wasn't. When ProPublica ran their analysis they discovered that there was a very high **False Positive Rate** for Black defendants compared to white defendants. Northpointe ran their own analysis a year later and tested for fairness using a **False Negative Rate**, coming to the conclusion that their algorithm was in fact fair.

From the context of the article and dispute that Northpointe and ProPublica had, what do you think a **False Positive** and a **False Negative Rate** means? Define False Positive rate and False Negative Rate in the context of this situation and explain how using each metric could lead to different conclusion about the fairness of the COMPAS algorithm.

LOAN DISCRIMINATION SIMULATION

Go to [this website](#) and **individually** read the context and instructions. Then, gather with your group and try toggling the different loan setting buttons.

If your group were building an AI to determine if someone should get a loan, which loan strategy would you choose and why? What metric of fairness would you use, if any? How did the fairness strategy you selected affect the bias of your model?

FINAL REFLECTION

What do you think is the relationship between bias and fairness? Write your name and thoughts below

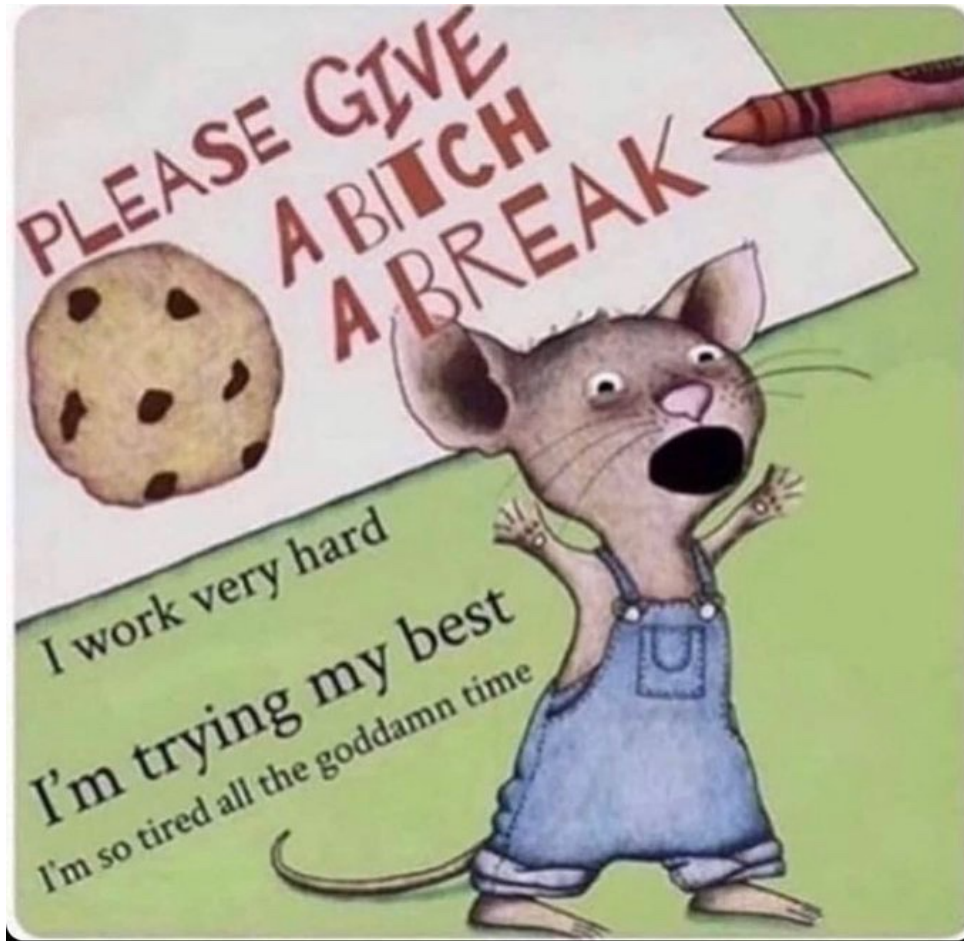
READING/WATCHING FOR TONIGHT

- [How to define fairness to detect and prevent discriminatory outcomes in Machine Learning](#)
 - [YouTube Weighs Major Changes to Kids' Content Amid FTC Probe](#)
-

REFLECTION (end-of-class)

Please take a moment to leave a note about something interesting you thought about or learned in class today and/or questions that the class raised for you. (Please include your name by default, but you are also welcome to include an anonymous note.)

Thursday 7/28 - Design Me!



Announcements:

-

In the News:

-

Agenda:

-

Additional Resources

Songs of the day

- Silent Hill - Kendrick Lamar
- Pride - Kendrick Lamar
- Chapter 6 - Kendrick Lamar
- Amor Prohibido - Selena

INDIVIDUAL REFLECTION

What is a recommender system that you've encountered? Write your name and answer below

WHAT'S FAIRNESS GOT TO DO WITH IT?

Spotify is an example of a multistakeholder recommendation system. Spotify, the company, must curate a list of music for users to listen to when they log on and in doing so they must balance the preferences of the musicians with those of the users. Musicians want to be discovered, users want to listen to music they like. Spotify has noticed that its Spanish speaking musicians have received a shockingly low proportion of listeners each month. Maybe this is because users just don't like Spanish music or they just aren't seeing any Spanish music recommendations.

Answer the following questions in your groups:

- (1) How should Spotify address the disproportionately low number of Spanish music listeners?
- (2) If Spotify starts to intentionally bias its recommendation system to push Spanish music more aggressively, would this be unfair? Why or why not?

YOUTUBE SCAVENGER HUNT!!

Work in your groups to uncover the ways in which YouTube uses AI. For each distinct feature, (1) describe the data that is being used to create the algorithm, (2) what the algorithm is trying to learn, (3) the overall "opinion" of the algorithm.

Name	Data	What is it trying to "Learn"?	What is it optimizing for?

YOUTUBE REDESIGN - PT1

In this activity, you will work with your group to redesign YouTube features according to various stakeholders' values. You will do this activity in two steps.

In step 1, you will identify 5 stakeholders (not including YouTube itself) and 2 values per stakeholder (e.g. things they care about).

In step 2, you will fill out the large paper matrixes on the walls around the classroom, filling in the boxes to mark which stakeholders care about which values. Stakeholders go in the first column, values go in the first row.

Complete step 1 with your groups in this google doc.

Stakeholder	Value #1	Value #2

YOUTUBE REDESIGN - PT2

Answer the following in your groups:

- (1) Given the ethical matrix above, identify the goal of your algorithm:
- (2) Why did you decide on this goal? Explain in a few sentences:
- (3) Given the goal of your algorithm that you decided above, please describe what data you need to teach your algorithm the correct actions:
- (4) What features will your re-designed YouTube platform have to achieve this goal?

Group 1

Group 2

Group 3

Group 4

READING/WATCHING FOR TONIGHT

- [Algorithms and Bias. Explained. Recode](#)

REFLECTION (end-of-class)

Please take a moment to leave a note about something interesting you thought about or learned in class today and/or questions that the class raised for you. (Please include your name by default, but you are also welcome to include an anonymous note.)