

Types of Biases in Data-driven Technology



Lecture 8

COMPAS → Predict whether a person is likely to offend again.

An 18-year-old African American girl was arrested in 2014 for the theft of a bicycle. She was charged with burglary amounting to \$80. More recently, a 41 years old Caucasian man was picked up for shoplifting tools worth \$86. The man was a repeat offender and had been previously convicted of many thefts and armed robberies. The girl had committed some minor offences when she was younger.

COMPAS identified the girl as High Risk & the Man as Low Risk



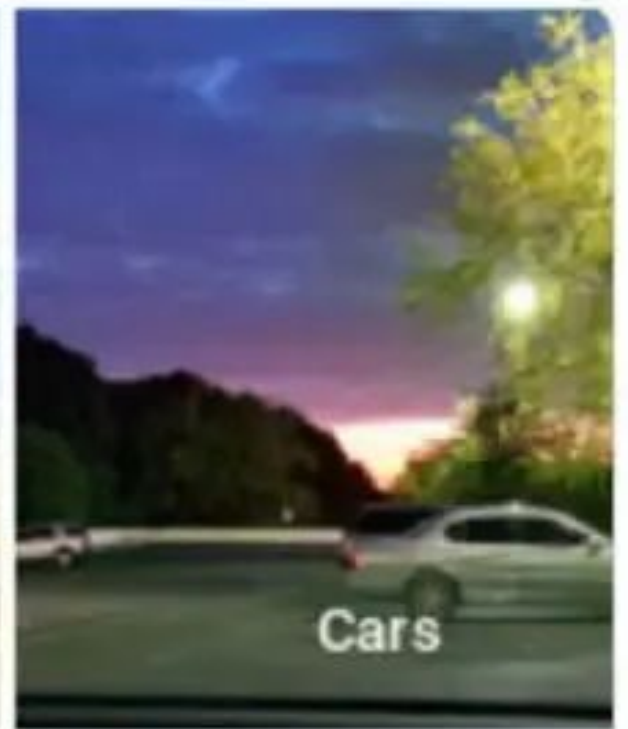
Google Photos



Skyscrapers



Airplanes



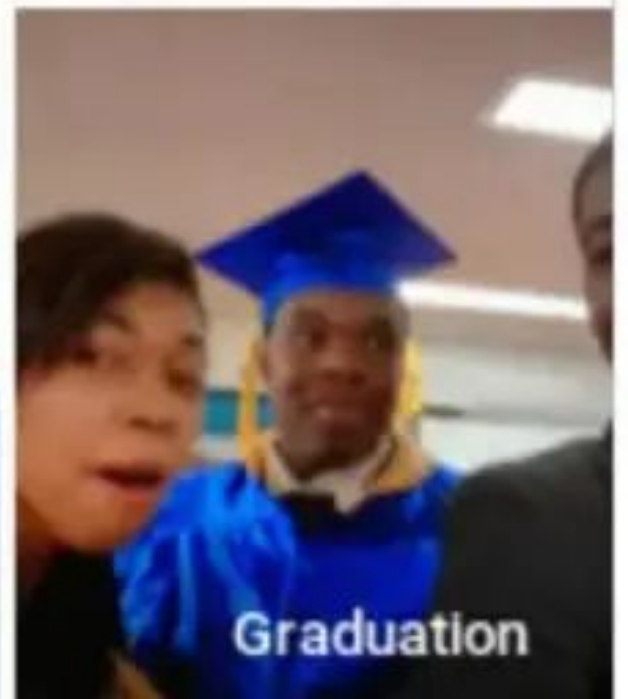
Cars



Bikes



Gorillas



Graduation

What is Bias?

Bias

Prejudice in favor of or against one thing, person, or group compared with another. That preference influences understanding and outcomes in a way that **prevents neutrality or objectivity**. Can be **positive**, but most often portrayed as **negative**.

Bias in Machine Learning

When the **predictions or decisions** obtained from different Machine Learning models are **unfair**.

Cause of Bias in Data-driven Technology

- Sometimes a dataset **maybe not as comprehensive** as it is needed to reflect reality.
- Sometimes **existing decisions or perceptions** that people have tagged information with are used as **samples by machines**.

Sociological Bias

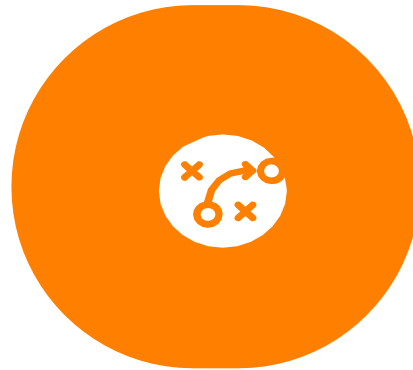
Racism (Ethnocentrism)



Americans: "British drive on the **wrong** side of the road rather than the **other** side."

Leads to racial profiling

Sexism



"Men are stronger and do all the work."

Leads to biased recruitments

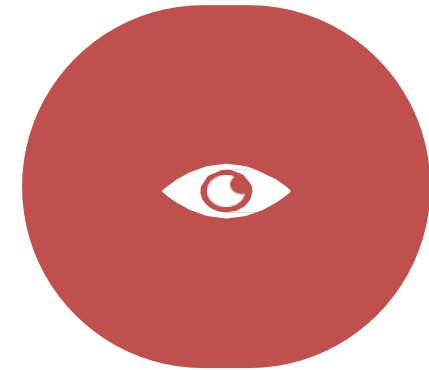
Ageism



"Older people are incapable of using techs"

Leads to social distances

Religious prejudice



"Muslims are terrorists and intolerant to other religions"

Leads to genocide, apartheid and many violent conflicts.

Implicit Bias

When we have **negative attitudes or stereotypes** towards a person, or group of individuals that we aren't even **consciously aware of**. - Google Translator

Bias in Society



Not accessible through introspection

May be the **opposite** of what we consciously think about **our beliefs**

Pervasive

Everyone possesses them

Infinity or Affinity Bias

Hiring someone with **similar interests** (doesn't help team **grow** or **diversify**)

Caused by exposure or lack of it

Exposure to stereotypes from friends and family or **lack of knowledge**

Implicit Bias

Impact on Society

They mainly predict **group actions**, especially when we are with a group of people who are like us.

It can lead to **systemic injustices** against people who are marginalized or treated unfairly in society.

For example, someone with implicit biases about race might not act in overtly racist ways, but their implicit bias still contributes to the passive support of any systemic racism when they're members of that majority group.

The issue with implicit bias is that people **are not aware** they are contributing to actions or systems that negatively impact individuals and groups.

How to prevent it?

Have an open mind and be willing to **unlearn**.

Increase awareness, embrace diversity, exposure to counter-stereotypical examples, and work together on a common task with equal status.

Project Implicit

Cognitive Bias

A systemic error in thinking that affects the decisions and judgments that people make when they are processing information about the world around them.



Anchoring

We tend to depend too heavily on a **new piece of information** offered than any subsequent information during decision-making

Example: Salary negotiations. The first number stated becomes the anchoring point for all negotiations.

Doctors diagnosing patients based on first impressions.



Fundamental attribution error

This describes how people will say their negative behavior was caused by a situation but for another person, it was a fundamental personality trait that caused the same negative behavior.

Student: "I was late because of traffic."

Teacher: "It's because you are lazy."

Cognitive Bias

A systemic error in thinking that affects the decisions and judgments that people make when they are processing information about the world around them.



Group attribution error

Refers to people's tendency to believe that the attributes of an **individual** **reflect the characteristics of the groups** they belong to.

Example, just because a team has made a decision, don't assume that everyone agrees.



False consensus effect

Describes the tendency to overestimate how normal our values and ideas are. It is strongest when we feel an issue is important, confident in our point of view, or convinced of the absolute moral rightness of the idea.

Example, online class.

Cognitive Bias

A systemic error in thinking that affects the decisions and judgments that people make when they are processing information about the world around them.



Optimism bias

It is the belief that our individual chances of experiencing something negative are lower than the people around us. The optimism bias makes us believe our chances of **experiencing something positive are higher than our peers.**

**Cognitive Bias can also allow
people to be manipulative**

Confirmation Bias

the tendency to analyze information in such a way that it **confirms your prior beliefs** while ignoring information that conflicts with it.

Bias in Thinking

Develop a Hypothesis



Scientists approach a research problem with a preconceived hypothesis

Search for favorable information



Scientists search for information favorable to their hypothesis & ignore other disproving data

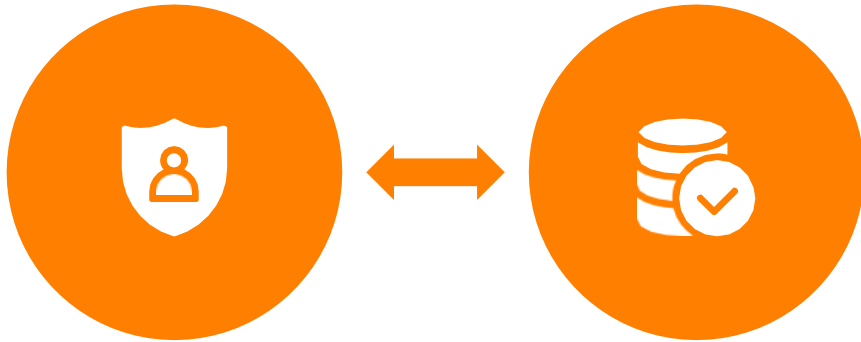
Publish the results



The fabricated results get published which supports their hypothesis

Complacency and Automated Bias

Automation Bias



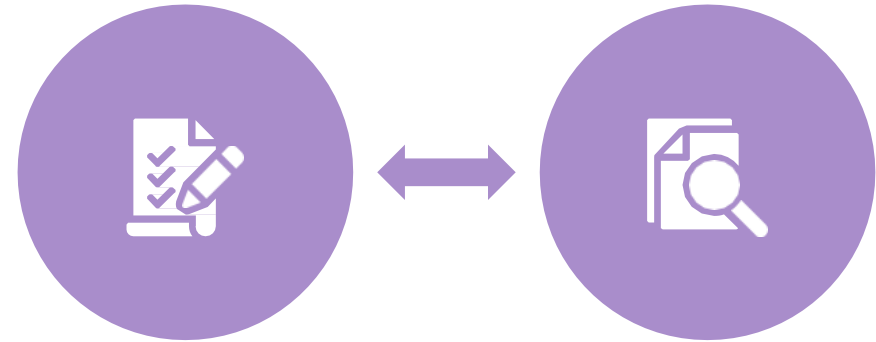
Trust

People trusts
automated system
over human decisions

Disregard

People disregards
search for contradictory
information

Complacency Bias



Reliable

People feel like
automated systems are
so reliable that they do
not require monitoring

Omission Errors

Machine fails to
detect a problem but
nobody notices

Example: Spell Check, GPS, Medicine

Data Collection Bias

“Garbage In, Garbage Out” - If you don't have good data, you won't get good results.

Bias in Data

Selection Bias

Not randomized data collection



Time interval Bias

Select a time period instead of a wide range



Observer Bias

People judge with cognitive bias and are not objective



Misclassification Bias

Incorrectly classified data



Selection Bias

Selecting everyone from a demographic for vote prediction

Time Interval Bias

Make estimates of profit on a single month's report where profit was high and cost was low

Observer Bias

Sending out a satisfactory survey right after conducting a workshop

Misclassification Bias

Where data is misclassified and rigs the result

Statistical Bias

01

Outliers

Outliers can be removed from dataset. If you had a set of scores of 25, 29, 32, 110, 33, 27, and 28, the 110 is an outlier.

02

Underfit

Removing variables might cause the ML model to underfit. As sometimes important information is left out.

03

Solution to overestimating

Include as many variables as possible before making a judgement of exclusion

04

Remove uncorrelated data

Remove data which are uncorrelated.
- Death by drowning is not related to ice creams sold!



4 Problems that biases help us address

Too much information

There is just too much information in the world, we have no choice but to filter almost all of it out.

01



Not enough meaning

We only see a tiny part of the world, but we need to make some sense of it in order to survive.

02



What to remember?

We can only afford to keep around the information that are most likely to prove useful in the future.

04



Need to act fast

We're constrained by time and information, and yet we can't let that paralyze us.

03



Home Work



Bias in Face Detection - Case study