Computer Science Ethics

15-110 – Monday 12/02

Announcements

- **HW6** due **Friday**!
- Fill out FCEs!!!

Learning Goals

- Understand the current extent of **data collection** on the internet and how data is used

- Understand the **notice-and-choice** model and its common criticisms

- Identify the societal impact when **automated decision-making** replaces human decision-making, including the effects of **bias**, **algorithmic appreciation**, **accountability**, and **explainability**

Ethics in Computer Science

Discuss: What does “ethics” mean to you?

Ethics in Computer Science

When we move from theoretical concepts of computer science to applying those theories in real life, the decisions we make have consequences.

The professional field of computer science has only recently adopted a code of ethics, and the code is not yet uniformly taught to new computer scientists or programmers. There is still much to debate over what the responsibilities of computer scientists are.

We'll discuss two areas where people debate how computing should be used in the current time: data collection and automated decision-making.

Data Collection

Many websites are funded by **advertising**, and consequently **data about consumers** is very valuable.

They provide content or functionality to consumers for free, and the cost of creating and maintaining the website is covered by advertisers, who pay to have the website show ads to the consumers.

Advertisers want to show ads only to users who are most likely to buy the product or service.

Therefore, data about consumers is valuable, because it lets advertisers direct ads most effectively

Advertisers want to create **profiles of consumers**: a collection of facts about them.

Profiles may include:

- Demographics: age, gender, race, etc
- Market participation: income level, location
- Preferences: family information, taste in books and movies, favorite restaurants or travel destinations, etc

Websites have a strong incentive to get the **best data possible** on their users, so they get paid more for advertisements.

This has led to **hyper-targeting** in ads, with ads attempting to reach more niche populations.

Check out the categories that you're hyper-targeted as belonging to:

- Facebook/Instagram: Settings and Privacy > Settings > Accounts Center > Ad Preferences > Ad Topics
- Twitter: Settings and Support > Settings and Privacy > Privacy and Safety > Ads Preferences > Interests
- TikTok: Settings and Privacy > Ads > How your ads are personalized

Discuss: what kinds of data are you comfortable having collected by companies? Where might you draw a line?

How are websites **collecting all this information** about users?

Everything we do reveals a little information:

- If I buy cat food, I probably have a cat
- If I look up the music video for a song, I probably like the song
- If I give my birthday when signing up for an online account, the website knows how old I am

This probably isn't too surprising: Spotify knows what music you like and the grocery store knows your favorite foods.

To be useful for advertising, though, advertisers need all of this information in one place, not scattered across lots of different companies.

When we talk about data collection and privacy, we're usually really talking about **data sharing**.

When a consumer interacts with a company, and a different company learns something about the interaction, we say that the other company is a **third party**.

This kind of data sharing is easy if both companies know a unique piece of information about the user, like the email address they used when creating an account. But people don't always make accounts, and sometimes they use different email addresses.

This means that the core of data sharing is **tracking**: finding a way to link together all the actions that a person takes on different websites and at different times.

Tracking often relies on **cookies**, a piece of data that a website asks a browser to remember, to identify the people visiting a website.

Here are a couple ways cookies are used:

1. You visit a website and select the “dark mode” view. The website asks your browser to remember `"darkMode: True"`. Next time you visit the site, your browser sends the cookie to the website, and the website shows you the dark mode view automatically.

Tracking often relies on **cookies**, a piece of data that a website asks a browser to remember, to identify the people visiting a website.

Here are a couple ways cookies are used:

2. You visit a shopping website and, behind the scenes, it assigns you an ID number. It asks your browser to remember `"id: 2486019552205"`. At the same time, it stores information about which products you looked at. It doesn't know your name or email address, so it uses the ID as the key. Weeks later, you visit the site again and enter your email address to place an order. Your browser sends the cookie, and the company links your email address to the products you looked at the first time you were on the site.

Tracking often relies on **cookies**, a piece of data that a website asks a browser to remember, to identify the people visiting a website.

Here are a couple ways cookies are used:

3. You visit a website that displays an ad. Behind the scenes, the ad company asks your browser to remember `"id: 3526323097732"`. At the same time, the ad company stores information about which website you're visiting, using the ID as the key. When you visit a different website that displays ads from the same company, your browser sends the cookie, and the ad company updates their records with the new information. If lots of websites show ads from the same ad company, they can see a lot of what you do online.

The last type of cookie is called a **third-party cookie.**

Fingerprinting

First party cookies are usually used for preferences, remembering that a user is logged in, and so on. Third party cookies are mostly used for tracking, so some browsers allow users to block third-party cookies. In response, data aggregators began using **fingerprinting** to track users across websites.

Fingerprinting is a technique that gathers lots of little pieces of information about the computer visiting a website. With enough little pieces, the website can uniquely identify the computer next time it visits.

How is this possible? Your browser makes lots of information about your computer visible to the websites you visit.

Behind the scenes, your **browser or phone/computer** is sending additional information to the services you use.

This is not done maliciously – services can put this information to good use. For example, knowing the size of your screen lets a website show you the mobile or desktop version of the site. However, you may be surprised by some of the data being shared.

Check out the data your browser shares here: https://webkay.robinlinus.com/

There are plugins you can install that limit the information your browser sends, but this may also limit functionality of websites.

Trackers are pieces of code created and published by data aggregators: companies that want to collect lots of data on lots of people.

Other companies use the tracker's code while building their website (like using an external module!). The tracker adds some functionality to the website, like letting the company see which pages of their website get the most traffic.

At the same time, the tracker reports back to the data aggregator that made it with information about the people visiting the website.

Trackers can also be included in emails.

Trackers use a combination of cookies and fingerprinting to identify the website's visitors.

Regulators around the world have tried a variety of methods to balance the desire of some consumers for privacy with the desires of companies.

In the US, a model called **notice and choice** has historically been dominant.

Under the notice and choice model, a company can legally use and share data about a consumer if:

- They **notify** the consumer, usually in a privacy policy or terms of service document, and
- The consumer **chooses** to agree, either by checking an "I agree" box or by continuing to use the website

Discussion:
Notice and Choice Effectiveness

Discuss: If the goal of regulation is to support consumer control over their data, does notice and choice do so?

Consider:

- When a website shows you a privacy policy or terms of service document, what do you do?
- What options do you have if you object to the data collection described in the document?

Notice and Choice Problems

Researchers (from CMU!) estimated the opportunity cost if everyone in the US fully read the privacy policies for the websites they used:

$781 billion

Since accepting the privacy policy is usually required to use the website, consumers often have to choose between allowing data sharing or not using the service at all.

Many governments are experimenting with ways of **protecting consumer data**.

In California, the California Consumer Privacy Act (CCPA) gives consumers rights to know about data collection and reject some collection.

In the European Union, the General Data Protection Regulation (GDPR) similarly restricts some data sharing, and gives consumers additional rights.

The US has no nationwide data privacy law, but legislators, regulators, and civil society groups have all shown interest in possible future legislation.

If you want to **protect your data online**, you have options!

Most browsers let you block cookies and can request that websites do not track you. You can also restrict permissions given to websites and applications on your devices.

You can check what kinds of trackers your browser stops and what your fingerprint looks like here: https://coveryourtracks.eff.org/

You can see what trackers and fingerprinting techniques a website is using by entering it here: https://themarkup.org/blacklight

One factor that fingerprinting uses is your IP address. You can hide your IP address from the websites you visit using a VPN. CMU has a VPN (though then CMU will know which websites you're accessing): https://www.cmu.edu/computing/services/endpoint/network-access/vpn/how-to/

Automated Decision Making

Automation Potential

There are potentially enormous benefits to be gained by using machine learning and artificial intelligence to accomplish tasks and solve problems.

Humans aren't always great at hard tasks, and lots of tasks take a lot of time. It's possible that machines do some tasks more consistently and save humans time as well.

However, we must keep in mind that there are potential downsides to these algorithms as well.

Training AI systems requires a **massive amount of labelled data**, and some of the ways that data is collected are unethical.

US based tech companies (like OpenAI and Meta) are paying vulnerable workers $2 and hour in Kenya to sift through horrific online content to train AI.

Training AI systems requires a massive amount of labelled data, and some of the ways that data is collected are unethical.

Tech companies making large learning models, ran out of English language text available on the Internet, but still needed more data and cut corners (ex: violating copyright laws) to collect it:

https://www.nytimes.com/2024/04/06/technology/tech-giants-harvest-data-artificial-intelligence.html

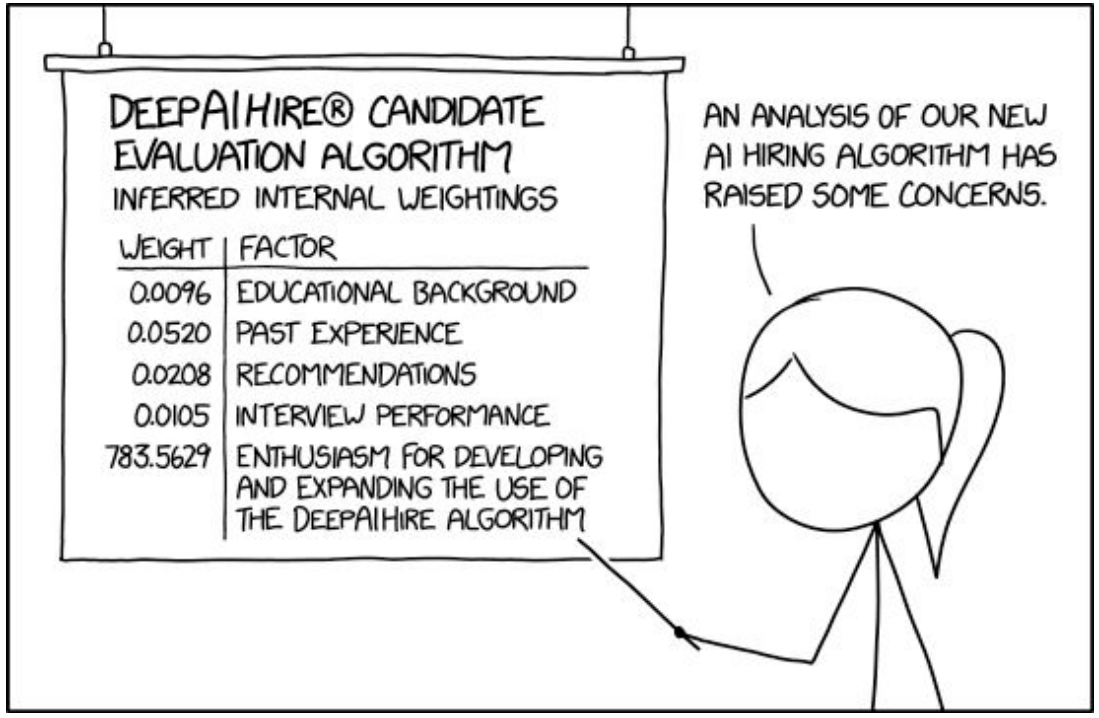
OpenAI created a voice assistant that sounded nearly identical to the voice of the actress Scarlett Johansson despite her clearly declining their offer to use her voice:

https://www.npr.org/2024/05/20/1252495087/openai-pulls-ai-voice-that-was-compared-to-scarlett-johansson-in-the-movie-her

Explainability

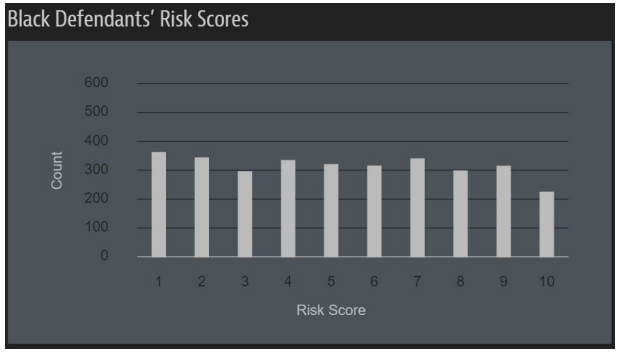
Decisions made by machine learning algorithms are usually based on a huge number of tiny features. In some algorithms (like neural networks) those features aren't named in a human-readable way.

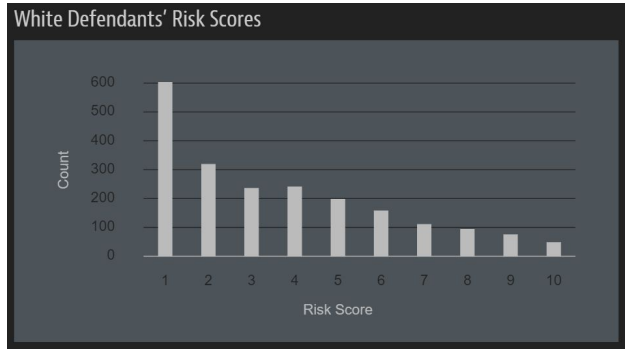
This is a problem when the algorithm makes an important decision about a person's life, like whether they should be admitted to a school or hired for a job.



Bias in the data fed into a machine learning algorithm or the design of the algorithm can lead to bias in the algorithm's results.

This has caused problems in algorithms for determining bail, which have shown systematic racial bias in predicting a person's likelihood to commit future crimes. This could be due to historical racial bias in bail decisions.





A similar problem was observed in an algorithm to hire engineers for Amazon, which showed bias towards hiring employees based on gender. Here the bias could be caused from the algorithm being trained on current employee resumes, where most of the current employees are male. Similar problems have been observed in other hiring algorithms too.

Prediction Fails Differently for Black Defendants

	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

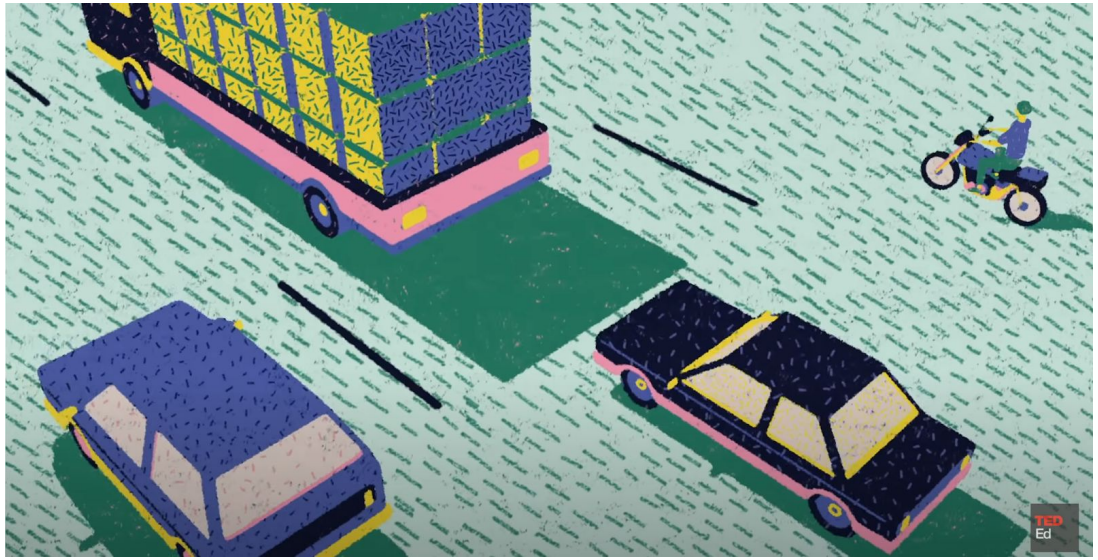
Overall, Northpointe's assessment tool correctly predicts recidivism 61 percent of the time. But blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend. It makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower risk but go on to commit other crimes. (Source: ProPublica analysis of data from Broward County, Fla.)

Ethics in AI design

Even if we set aside the problems related to bias in data (which obviously affect human decision making as well), there are still big ethical questions about how we should use AIs.

If an algorithm makes a mistake (say, a self-driving car crashes, or police use a facial recognition algorithm to identify a suspect, and it identifies an innocent person), it isn't always clear who should be held accountable.

Discuss: when should algorithm creators be held responsible for the unintended outcomes of their algorithms?



Responsibility Assignment

Questions about responsibility extend to smaller day-to-day actions algorithms may take too.

For example, Google has become a gatekeeper for much of the information in the world. If a small change to Google's search algorithm moves a small business from the first page to the second, that could have a drastic effect on the business's revenue.

This also applies to the algorithms social media networks use to decide which posts should be promoted. Studies have shown these algorithms can lead to the spread of false information.

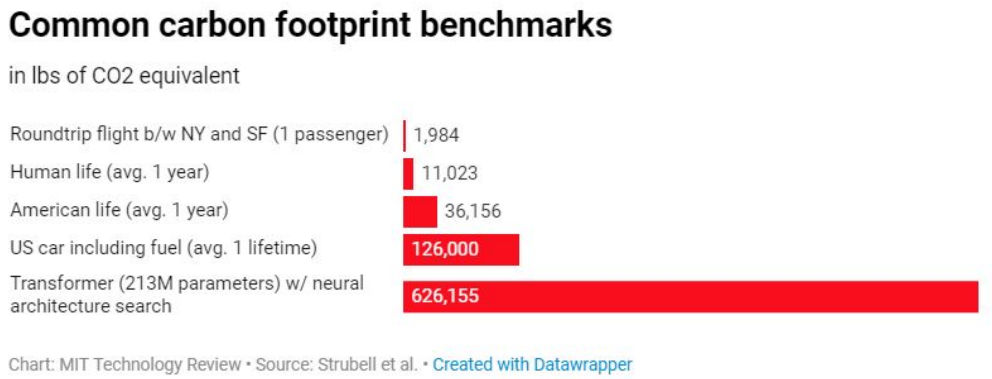
Sidebar: Effects on the Environment

Even when we make productive and unbiased algorithms, they can still have unintended side effects.

Many companies and researchers train machine learning algorithms on very large datasets to answer questions. This analysis does not come without a cost.

An enormous amount of energy is needed to run these algorithms, and in the US, that energy often has a carbon footprint. A recent study found that training a popular NLP model, The Transformer, left a gigantic carbon footprint.

On the bright side, some tech companies have pledged to go carbon negative to combat this. Other scientists are exploring new ways to make algorithms more energy efficient.



Learning Goals

- Understand the current extent of **data collection** on the internet and how data is used

- Understand the **notice-and-choice** model and its common criticisms

- Identify the societal impact when **automated decision-making** replaces human decision-making, including the effects of **bias**, **algorithmic appreciation**, **accountability**, and **explainability**