IS4246
Smart Systems and AI Governance

Lecture 5

NUS
National University of Singapore
National University of Singapore

Agenda

- Review From Last Time

- Privacy in AI/ML
  - What's the big deal?
  - Issues with big data
  - What can we do?

Review from Last Time

Questions You Had From Last Week

What’s the Big Deal?

Companies use data collection and analysis to gain insights into customers' shopping habits and preferences.

Target Case

Video: https://shorturl.at/cgnwC

1.00
NATE SILVER ON WHAT OBAMA SHOULD DO NEXT, P. 44
DON'T MOCK THE ARTISANAL PICKLE MAKERS, P. 14
A NANNY'S VIEW OF THE WORLD, P. 47
MANAGEMENT TIPS FROM 'DOWNTON ABBEY,' P. 52
GREECE CONFRONTS ITS SPARTAN FUTURE, P. 38
The New York Times Magazine
February 19, 2012
HEY YOU'RE HAVING
RETAILERS PROFIT FROM PREDICTIONS
How your shopping habits reveal even the most personal information. By Charles Duhigg

Charles Duhigg, How Companies Learn Your Secrets, N.Y. Times Magazine, Feb. 16, 2012

Target Case

There are some brief periods in a person’s life when old routines fall apart and buying habits are suddenly in flux. One of those moments — *the* moment, really — is right around the birth of a child, when parents are exhausted and overwhelmed and their shopping patterns and brand loyalties are up for grabs.

“We knew that if we could identify them in their second trimester, there’s a good chance we could capture them for years. As soon as we get them buying diapers from us, they’re going to start buying everything else too.”

Charles Duhigg, How Companies Learn Your Secrets, N.Y. Times Magazine, Feb. 16, 2012

Target Case

Target assigns each customer a unique number, known as a Guest ID, that tracks their purchasing habits. Connected to that Guest ID is information about the shopper, such as age, marital status, home location, estimated salary, credit cards owned, web activity, ethnicity, employment data, magazine subscriptions, bankruptcy history, house purchase or sale, and college attended.

Charles Duhigg, How Companies Learn Your Secrets, N.Y. Times Magazine, Feb. 16, 2012

Target Case

By analyzing data on shoppers' purchases, demographic information, and online activity, Target can predict and influence consumer behavior.

Understanding and manipulating habits can lead to transformative changes in organizations and societies, from improving worker safety to influencing voting patterns.

Charles Duhigg, How Companies Learn Your Secrets, N.Y. Times Magazine, Feb. 16, 2012

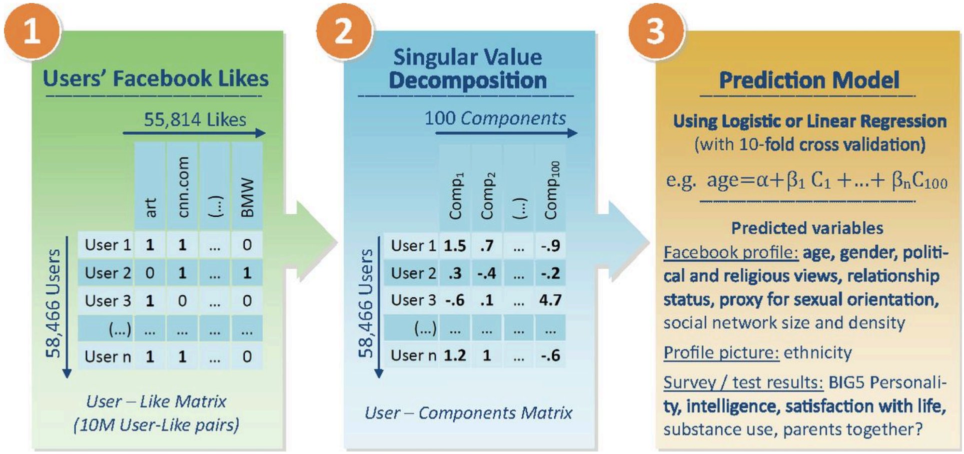
Target Case

- How does Target analyze customer data to make inferences about customers?

- In your opinion, is the Target system an intrusion on privacy? Why or why not?

Predicting Private Traits

- Facebook Likes can be used to accurately predict personal attributes such as sexual orientation, ethnicity, political views, and personality traits.



1

Users' Facebook Likes

55,814 Likes

58,466 Users

	art	cnn.com	(...)	BMW
User 1	1	1	...	0
User 2	0	1	...	1
User 3	1	0	...	0
(...)	...	...	...	...
User n	1	1	...	0

User – Like Matrix
(10M User-Like pairs)

2

Singular Value Decomposition

100 Components

58,466 Users

	$Comp_1$	$Comp_2$	(...)	$Comp_{100}$
User 1	1.5	.7	...	-.9
User 2	.3	-.4	...	-.2
User 3	-.6	.1	...	4.7
(...)	...	...	...	...
User n	1.2	1	...	-.6

User – Components Matrix

3

Prediction Model

Using Logistic or Linear Regression
(with 10-**fold cross validation)**

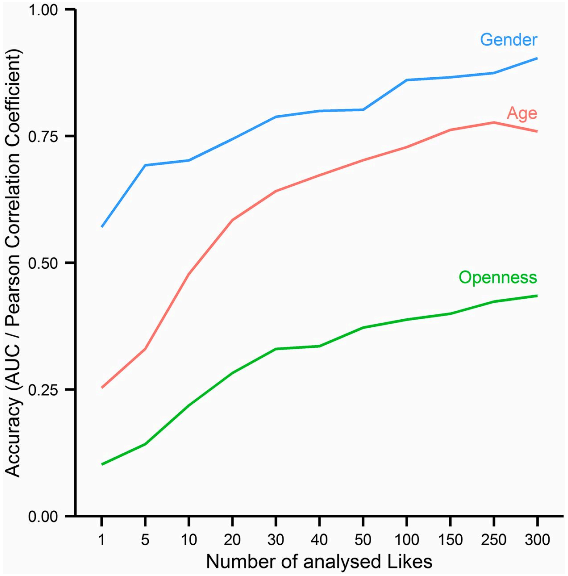
e.g. $age=\alpha+\beta_1 C_1 +\ldots+ \beta_n C_{100}$

Predicted variables

Facebook profile: **age, gender, political and religious views, relationship status, proxy for sexual orientation,** social network size and density

Profile picture: ethnicity

Survey / test results: BIG5 Personality, **intelligence, satisfaction with life,** substance use, parents together?



Michal Kosinski et al, Private Traits and Attributes are Predictable from Digital Records of Human Behavior, Proceedings of the National Academy of Sciences, April 9, 2013

Predicting Private Traits

- Attributes can be inferred from other digital records, including browsing histories, search queries, and purchase histories

- Commercial companies, governmental institutions, or even one's Facebook friends could use software to infer attributes such as intelligence, sexual orientation, or political views that an individual may not have intended to share

Michal Kosinski et al, Private Traits and Attributes are Predictable from Digital Records of Human Behavior, Proceedings of the National Academy of Sciences, April 9, 2013

Predicting Private Traits

- What could possibly go wrong?

Michal Kosinski et al, Private Traits and Attributes are Predictable from Digital Records of Human Behavior, Proceedings of the National Academy of Sciences, April 9, 2013

The Role of Consent and Anonymity

- Privacy protections for the past 40 years have concentrated on two types of procedural mitigations: informed consent and anonymization.
- Over time researchers have documented serious cracks in the protections afforded by informed consent and anonymity.
- Long-standing operational challenges come to a head with Online Behavioral Advertising.

Barocas, S., & Nissenbaum, H. (2014). Big data's end run around procedural privacy protections. *Communications of the ACM*, *57*(11), 31-33.

The Inevitability of Big Data

- Big data extinguishes what little hope remains for the you-notice and choice regime.
- Weaknesses in existing procedures do not undercut the validity of privacy itself.

Barocas, S., & Nissenbaum, H. (2014). Big data's end run around procedural privacy protections. *Communications of the ACM*, *57*(11), 31-33.

Anonymity 2.0

- Most online outfits claim to keep anonymous records, yet they still employ unique *persistent identifiers* – IDs or pseudonyms
- This limits the likelihood of leaked *Personal Identifying Information (PII)*
  - Any type of data that can be used to identify someone, from their name and address to their phone number, passport information, and social security numbers
- Re-identification still very possible

Barocas, S., & Nissenbaum, H. (2014). Big data's end run around procedural privacy protections. *Communications of the ACM*, *57*(11), 31-33.

Justice and Beneficence

- In biomedicine, informed consent and anonymity are not the only protective mechanisms in play.
- Treatment or research protocols must also pass tests of *justice* and *beneficence*, and consent forms must go through ethical scrutiny via institutional review boards (IRB).
- Consent forms have undergone ethical scrutiny and been thoroughly debated, with the individual's signature enacting welfare values

Barocas, S., & Nissenbaum, H. (2014). Big data's end run around procedural privacy protections. *Communications of the ACM*, *57*(11), 31-33.

Justice and Beneficence

- Why may informed consent be insufficient?

Barocas, S., & Nissenbaum, H. (2014). Big data's end run around procedural privacy protections. *Communications of the ACM*, *57*(11), 31-33.

Key Elements of (US) Privacy Law

- **The Right of Privacy**: The right of privacy is enshrined in the Fourth Amendment to the US Constitution. It protects citizens and permanent residents from unreasonable searches and seizures within their home, papers, or effects.

- **The Fair Credit Reporting Act (FCRA)**: The FCRA is a federal law that regulates what information companies can report about consumers' credit history and how they make those reports. It's designed to protect consumers from identity theft and fraud and promote accuracy in credit reporting.

Key Elements of (US) Privacy Law

- **The Health Insurance Portability and Accountability Act (HIPAA)**: This federal law sets restrictions on how healthcare providers, health plans, and any other organization that uses personal health information can use and share that information.

- **The Children's Online Privacy Protection Act (COPPA)**: COPPA is a law that requires websites, hosting services, and online applications that collect personal information from kids younger than 13 years old to get parental consent before collecting and sharing the data.

Discussion

- Do you think the U.S. legal framework is effective in general?

- Will it be effective for the novel challenges of big data and business analytics?

Feature Selection, Proxies, Masking

- “*Protected Classes*”: Information that the designers should NOT be used for data mining – e.g. race, gender, sexual orientation, etc.

- *Feature selection* decisions can have serious implications for the treatment of protected classes if those factors that better account for pertinent statistical variation among members of a protected class are not well represented.

“Rational Racism”

- *Rational racism* is when race is taken into consideration explicitly due to lack of access to other information.

- This is often an inaccurate assessment as race is too coarse an indicator to predict an individual’s outcome

Inadequacy of Representation:

- Data are reductive representations that may fail to capture enough detail to allow for the discovery of crucial points of contrast.

- Mapping out these mechanisms may be ineffective as they may not lend themselves to exhaustive representation in the data.

- What can be done to avoid using or impacting these protected classes adversely?

- What can be done to avoid using or impacting these protected classes adversely?

- Remove the “protected features”?

Issue 1: Proxies

- *Unintentional discrimination* can occur when criteria that are relevant to decision making also serve as proxies for class membership.
- "*Redundant encodings*" of class membership in other-data-pieces can facilitate this.
- Better data and more features may further expose the inequality that already exists.

Issue 1: Proxies

- Example 1: Paying attention to the most competent employees may inadvertently disadvantage protected classes.

Issue 1: Proxies

- Example 2: Data mining to maximize accuracy can lead to disparate impacts.

Issue 2: Masking Intentional Discrimination

- Data mining can mask intentions of intentional discrimination
- Protected classes may be disadvantaged without direct access to information

Suggestion:

- Focus on ‘unintentional discrimination’ as it is more commonly easier to overlook

Differential Privacy

- The U.S. Census Bureau uses mathematical concept called differential privacy to release 2020 census data
- Census Bureau’s job to collect, analyze, and disseminate useful information
- At the same time, the bureau is prohibited by law from releasing any information for which "the data furnished by any particular establishment or individual … can be identified."
- Video: https://www.youtube.com/watch?v=gI0wk1CXlsQ&ab_channel=SimplyExplained

Jeffrey Mervis, Can a Set of Equations Keep U.S. Census Data Private?, Science, January 4, 2019

Differential Privacy

- *Past Solutions*
  - Removing names and addresses
  - Injecting noise
- Issue: the Data Reconstruction Theorem
  - Theorem shows that, given access to a sufficiently large amount of information, someone can reconstruct underlying databases and, in theory, identify individuals.
  - Each query consumes a little bit of what the experts call a "privacy budget." After that budget is exhausted, queries are halted in order to prevent database reconstruction.
  - In the case of census data, however, the agency has already decided what information it will release, and the number of queries is unlimited. So its challenge is to calculate how much the data must be perturbed to prevent reconstruction.

Jeffrey Mervis, Can a Set of Equations Keep U.S. Census Data Private?, Science, January 4, 2019

Differential Privacy

- *Issues*
  - Increases cost for surveys, taking more time
  - Social scientists get worse information
  - Potentially not needed of the data “ages” over time and you can’t use past years to “follow” individuals

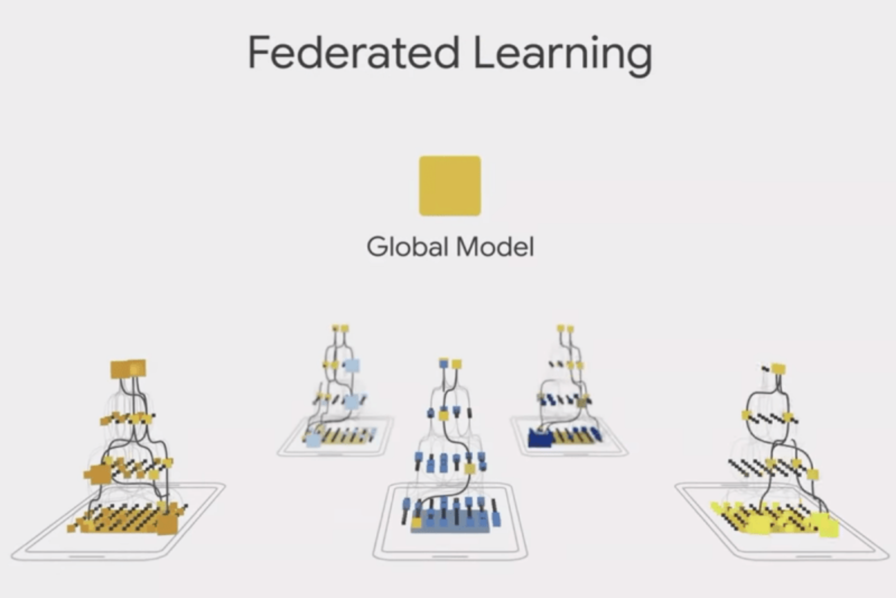
Jeffrey Mervis, Can a Set of Equations Keep U.S. Census Data Private?, Science, January 4, 2019

Federated Learning

- https://youtu.be/zqv1eELa7fs?si=VYzLWygVgMym8eq_

- *Federated learning* enhances privacy by training AI models on mobile devices without the need for data to leave the device.

- Federated learning offers advantages such as improved *privacy*, reduced *communication rounds*, and the ability to *analyze protected data*, making it suitable for industries like healthcare.



Khari Johnson, How Federated Learning Could Shape the Future of AI in a Privacy-Obsessed World, Venturebeat, August 9, 2019 • What is federated learning, and how could it address privacy concerns around analytics?

- What challenges or opportunities do you see around Federated Learning?

Thank You!