

# Data Science and Ethics

Giri Iyengar

Cornell University

*gi43@cornell.edu*

May 2, 2018

# Agenda for the day

- Privacy, Ethics, and Data Science
- Course Summary
- Q&A

# Overview

1 Privacy, Ethics, and Data Science

2 Course Summary

3 Questions and Answers

# How Data Science can violate Privacy

- Target's use of Data Mining raises concern
- Science: The End of Privacy
- Is your Browser Safe?
- Lack/Too much Privacy threatens Democracy

# Problems with violation of Privacy

- Information used out of context
- Unfair treatment (e.g. medical insurance, employment)
- Identity Theft
- Unique problems posed by the Internet (it doesn't forget!)

## Automated Prediction Systems

A non-interpretable process might follow from a data-mining analysis which is not explainable in human language. Here, the software makes its selection decisions based upon multiple variables (even thousands). It would be difficult for the government to provide a detailed response when asked why an individual was singled out to receive differentiated treatment by an automated recommendation system. The most the government could say is that this is what the algorithm found based on previous cases.

# Perpetuation of Biases

- Can Computers be Racist?
- Suggestive Ads
- Strengthening human bias
- Truly good intentions gone awry

# Amplification of Privilege

- How we amplify privilege with supervised machine learning
- Strata presentation on Amplification of Privilege

# Acceptable False Alarm Rate?

- SKYNET program



# Filter Bubble

- The Filter Bubble
- What the Internet is Hiding from you

# Overview

1 Privacy, Ethics, and Data Science

2 Course Summary

3 Questions and Answers

# Summary

- What is Data Science?
- Data Processing Tools - Map Reduce / Spark
- Machine Learning
- Recommendation Systems
- Graph Analysis
- Deep Learning
- Data Structures, Sampling, Streaming techniques

# Overview

- 1 Privacy, Ethics, and Data Science
- 2 Course Summary
- 3 Questions and Answers

## Questions and Answers

