Bradley Voytek, Ph.D.

UC San Diego

Department of Cognitive Science
Halıcıoğlu Data Science Institute
Neurosciences Graduate Program

bvoytek@ucsd.edu
@bradleyvoytek

5 Assignments (10% each)

Assignments are completed individually and graded programmatically.

Instructions must be followed to receive credit.

These are meant to get you practice programming around the topics covered in class.

The first one is much simpler than the following four and will take less time.

Assignments will be due on Sundays by 11:59 PM

Assignments @ Datahub

DATA SCIENCE / MACHINE LEARNING PLATFORM

UC San Diego

UC San Diego Jupyterhub (Data Science) Platform

Project proposal (10%)

Project Proposals are Completed in Assigned Groups of 4-6.

These are due Sunday (11:59 PM) after Week 3.

These include a proposal of what question you are setting out to answer for your final project and what data you plan to use to complete this project.

Final project (35%)

The final project is a Jupyter notebook completed with your assigned group carrying out a complicated data science project from start to finish.

This should be completed throughout the quarter. Being a contributing member of a team is a requirement of this project. You should start on this as early as possible and work on it a little each week.

Data science ethics

LIBERAL-ARTS MAJORS MAY BE ANNOYING SOMETIMES, BUT THERE'S *NOTHING* MORE OBNOXIOUS THAN A PHYSICIST FIRST ENCOUNTERING A NEW SUBJECT.

Source: *XKCD*

Data science ethics

"Universities should start with broader training for computer science students… Such courses are hard to teach well. Computer scientists often don't take them seriously, are uncomfortable with non-quantitative thinking, are overconfident because they're mathematically brilliant, or are convinced that utilitarianism is the answer to everything.

But universities need to try. Professors need to scare their students, ***to make them feel they've been given the skills not just to get rich but to wreck lives***; they need to humble them, to make them realize that however good they might be at math, there's still so much they don't know."

Source: Emma Pierson, "Hey, Computer Scientists! Stop Hating on the Humanities", *Wired*

Big data ethics

"Big data and analytics technology can reap huge benefits to both individuals and organizations—bringing personalized service, detection of fraud and abuse, efficient use of resources and prevention of failure or accident. **So why are there questions being raised about the ethics [of data science]?**"

Source: IBM

YouTube vows to recommend fewer conspiracy theory videos

Site's move comes amid continuing pressure over i platform for misinformation and extremism

The Reason This "Racist Soap Dispenser" Doesn't Work on Black Skin

Amazon Prime and the racist algorithms

MACHINES TAUGHT BY PHOTOS LEARN A SEXIST VIEW OF WOMEN

Facial recognition software is biased towards white men, researcher finds

Biases are seeping into software

YouTube's Restricted Mode Is Hiding Some LGBT Content [Update]

Google Translate's Gender Problem (And Bing Translate's, And Systran's...)

Adapted from Thomas Donoghue

Always consider ethics.

ETHICS

"Moral principles that govern a person's behaviour or the conducting of an activity."

Source: Oxford Dictionary

UC San Diego

Big data ethics

Source: IBM

Ethical data science

Data science pursued in a manner so that is equitable, with respect for privacy and consent, so as to ensure that it does not cause undue harm.

INTENT and OBJECTIVITY

- Intent is not required for harmful practices to occur.

- Data, algorithms and analysis are not objective.
    - It is done by people, who have biases.
    - It uses data, which have biases.

- Data Science is powerful

- Bias & discrimination driven by data & algorithms can give new scale to pre-existing inequities.

On the law

- There are laws that cover research, privacy and other aspects of data use, data science, and algorithms
- We're not really talking about the law today
- Regulations are needed

NINE THINGS TO CONSIDER TO NOT RUIN PEOPLE'S LIVES WITH DATA SCIENCE

Adapted from Thomas Donoghue

UC San Diego

1. THE QUESTION

- What is your question? Is it well-posed?
- Do you know something about the context and background of your question?
- What is the scope your investigation? What correlates might you inadvertently track? Is it possible to answer this question well?

Case Study: Labelling Faces

Detecting criminality from faces [link, paper]

(a) Three samples in criminal ID photo set S_c.

(b) Three samples in non-criminal ID photo set S_n

Figure 1. Sample ID photos in our data set.

Detecting Sexual Orientation From Faces with computer vision [link, paper]

2. THE IMPLICATIONS

- Who are the stakeholders? How does this affect them?
- Could the information you will gain and/or the tool you are building be co-opted for nefarious purposes?
    - If so, can you protect them from that?
- Have you considered potential unintended consequences?

Case Study: Abuse of social networks

The New York Times

A Genocide Incited on Facebook, With Posts From Myanmar's Military

Facebook has been co-opted by military personnel to spread misinformation, hate speech, and promote ethnic cleansing [news link, UN Report]

3. THE DATA

- Is there data available? Is this data directly related to your question, or only potentially related through proxies?
- Who do you have data from?
- Do you have enough data to make reliable inferences?
- What biases does your data have?
- If you do not have, and can not get, enough good, appropriate data, you may just have to stop.

Case Study: Biomedical Science

Biomedical research has often excluded female subjects

This was based on a (faulty) assumption that females would be more variable

These findings do not generalize as well

Sources: link, link, link

ASIDE: RESEARCH ETHICS

RESEARCH: A systematic investigation, including research development, testing, and evaluation, designed to develop or contribute to generalizable knowledge.

Investigations that qualify as research have their own regulatory structure that must be followed

The Belmont Report

The Belmont Report, published in 1974 by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, is a statement of **basic ethical principles and guidelines** that should assist in resolving the ethical problems that surround the conduct of research with human subjects.

Source: Collaborative Institutional Training Initiative

The Belmont Report

The Belmont Report presents three principles that today serve as basic concepts in the conduct of research with human subjects:

- **Respect for Persons**
- **Beneficence**
- **Justice**

The principles are intended to "assist scientists, subjects, reviewers and interested citizens to understand the ethical issues inherent in research involving human subjects." It is generally accepted that the principles share equal weight, and the evaluation and conduct of human subjects research are done in the context of all three.

Source: Collaborative Institutional Training Initiative

The Common Rule

The Federal Policy for the Protection of Human Subjects

- Requirements for assuring compliance by research institutions
- Requirements for researchers' obtaining and documenting informed consent
- Requirements for Institutional Review Board (IRB) membership, function, operations, review of research, and record keeping.
- Protections for certain vulnerable research subjects
  - pregnant women
  - *in vitro* fertilization and fetuses
  - prisoners
  - children

Beneficence and risk/benefit analysis

"Persons are treated in an ethical manner not only by respecting their decisions and protecting them from harm, but also by **making efforts to secure their well-being**."

"Beneficence… requires that we protect against risk of harm to subjects and also that we be **concerned about the loss of the substantial benefits that might be gained from research**."

Source: Michelle Meyer

Beneficence and risk/benefit analysis

"Research also makes it possible to avoid the harm that may result from the application of previously accepted routine practices that on closer investigation turn out to be dangerous."

"[E]stimates of the probability of harm [must be] reasonable, as judged by known facts or other available studies."

Source: Michelle Meyer

Respect for persons and informed consent

Don't "withhold information necessary to make a considered judgment, **when there are no compelling reasons to do so**."

"In **most** cases of research involving human subjects, respect for persons demands that subjects enter into the research voluntarily and with adequate information. In some situations, however, **application of the principle is not obvious**."

Source: Michelle Meyer

Respect for persons and informed consent

"A special problem of consent arises where informing subjects of some pertinent aspect of the research is likely to impair the validity of the research. In many cases, it is sufficient to indicate to subjects that they are being invited to participate in research of which some features will not be revealed until the research is concluded."

Source: Michelle Meyer

4. INFORMED CONSENT

INFORMED CONSENT: the voluntary agreement to participate in research, in which the subject has an understanding of the research and its risks

Informed consent can be withdrawn at any point in time

Case Study: Emotional Contagion

Facebook conducted an experiment investigating whether they could manipulate people's emotions by manipulating the content displayed on one's newsfeed. [link, paper]

The Facebook Study

- Over 1B served/mo worldwide (in U.S., 128m users/day)
- Average user: ~1500 items eligible for news feed
- Algorithm selects & shows ~300 items (to see all, go to friend's wall)
- Has always existed
- FB constantly changes it
- Proprietary
- Now based on ~100K criteria
- Almost certainly affects how much emotionally charged content users see; very likely prioritizes ☺️ posts

McGee M, "EdgeRank Is Dead: Facebook's News Feed Algorithm Now Has Close To 100K Weight Factors," *Marketing Land*, Aug 2013
http://marketingland.com/edgerank-is-dead-facebooks-news-feed-algorithm-now-has-close-to-100k-weight-factors-55908

Source: Michelle Meyer

UC San Diego

The Facebook Study: Three hypotheses

1. Social Comparison (☺️ posts are risky)
   - Small, observational studies
   - Correlations b/w FB use and stress, jealously, loneliness, and depression
   - FB creates "self promotion-envy spiral"

2. Emotional Contagion (☹️ posts are risky)
   - Lab experiments and field data found happiness and depression spread via in-person social networks

3. Null Hypothesis

Source: Michelle Meyer; Fowler & Christakis, "Dynamic spread of happiness in a large social network," *BMJ* 2008

The Facebook Study: Design

- 1 week (Jan. 11-18, 2012)
- Subjects "who viewed FB in English" randomly selected by user ID
- 2 experiments manipulated extent to which subjects (N = 689,003) were exposed to emotional content by *removing* varying %'s of it from particular news feed viewings
- All posts remained on friends' walls; post not presented in 1 viewing may appear in other viewings
- Software coded status updates as ☺️, ☹️, or neutral ("no text was seen by the researchers")
  - Coded +/- if at least 1 +/- word, e.g., "nice" or "awful"

Source: Michelle Meyer; Kramer, Guillory, Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *PNAS* 2014

The Facebook Study: Data distribution

Source: Michelle Meyer; Kramer, Guillory, Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *PNAS* 2014

The Facebook Study: Design

Two parallel experiments; 4 conditions (for each, N = ~155,000)

Experiment 1: Removing ☺️ posts

- Treatment: Each ☺️ post "had between a 10% & 90% chance (based on user ID) of being omitted…for that specific viewing."
- Control: 10%–90% of 46.8% (i.e., 4.68%-42.12%) of eligible posts randomly removed w/o regard to emotional content

Experiment 2: Removing ☹️ posts:

- Treatment: Each ☹️ post "had between a 10% & 90% chance (based on user ID) of being omitted…for that specific viewing."
- Control: 10%–90% of 22.4% (i.e., 2.24%-20.16%) of eligible posts randomly removed w/o regard to emotional content

Source: Michelle Meyer; Kramer, Guillory, Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *PNAS* 2014

The Facebook Study: Design

Two parallel experiments; 4 conditions (for each, N = ~155,000)

Experiment 1: Removing ☺️ posts

- Treatment: Each ☺️ post "had between a 10% & 90% chance (based on user ID) of being omitted…for that specific viewing."
- Control: 10%–90% of 46.8% (i.e., 4.68%-42.12%) of eligible posts randomly removed w/o regard to emotional content

Experiment 2: Removing ☹️ posts:

- Treatment: Each ☹️ post "had between a 10% & 90% chance (based on user ID) of being omitted…for that specific viewing."
- Control: 10%–90% of 22.4% (i.e., 2.24%-20.16%) of eligible posts randomly removed w/o regard to emotional content

Source: Michelle Meyer; Kramer, Guillory, Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *PNAS* 2014

The Facebook Study: Results

Compared to control subjects, subjects exposed to fewer ☺️ posts subsequently in their own posts:

- Used 0.1% fewer ☺️ words (Cohen's d = 0.02)
- Used 0.04% more ☹️ words (Cohen's d = 0.001)
- Produced only 96.7% as many words overall

Compared to control subjects, subjects exposed to fewer ☹️ posts subsequently in their own posts:

- Used 0.07% fewer ☹️ words (Cohen's d = 0.02)
- Used 0.06% more ☺️ words (Cohen's d = 0.008)
- Produced only 99.7% as many words overall

Source: Michelle Meyer; Kramer, Guillory, Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *PNAS* 2014

The Facebook Study: Results

Issues of methodology & interpretation:

- Questionable instrument: LIWC 2007 not intended for lengthy text
- Questionable coding: "I'm not having a great day." "Oh great."
- Effect on word choice ≠ effect on mood (no giggling at a funeral)

Source: Michelle Meyer

The Facebook Study: Results

"First, these effects, while highly statistically significant, are tiny. The largest effect size reported had a Cohen's d of 0.02–meaning that eliminating a substantial proportion of emotional content from a user's feed had the monumental effect of shifting that user's own emotional word use by two hundredths of a standard deviation."

Source: Tal Yarkoni

UC San Diego

The Facebook Study: Results

"To put it in intuitive terms, the effect of condition in the Facebook study is roughly comparable to a hypothetical treatment that increased the average height of the male population in the United States by about one twentieth of an inch (given a standard deviation of ~2.8 inches). Theoretically interesting, perhaps, but not very meaningful in practice."

Source: Tal Yarkoni

The Facebook Study: Objections

- No IRB review
- No informed consent (except ToS)
- No debriefing
- Knowingly psychologically harmed users
  - *Guardian*: "deliberately made people sad"
  - *Slate*: "intentionally made thousands upon thousands of people sad"

Source: Michelle Meyer

EULA

The new EULA reads: *"The non-personally identifiable information that EA collects includes technical and related information that identifies your computer (including the Internet Protocol Address) and operating system, as well as information about your Application usage (including but not limited to successful installation and/or removal), software, software usage and peripheral hardware."*

Source: gamesradar

EULA - Options?

that data with our third party service providers in a form that does not personally identify you. IF YOU DO NOT WANT EA TO COLLECT, USE, STORE, TRANSMIT OR DISPLAY THE DATA DESCRIBED IN THIS SECTION, PLEASE DO NOT INSTALL OR USE THE APPLICATION. This and all other data provided to EA and/or collected by EA in

Above: Either you agree to have your privacy invaded, or you can't play Origin games

Source: gamesradar