

DataVizA Tutorial: Ethics in Data Science

Department of Econometrics and Business Statistics, Monash University

Tutorial 4

Background

This week we're going to be looking into some case studies of ethical practice. Your tutor will split you into groups of 3-5 in different break out rooms through Zoom.

Each group will be given a short (paragraph) vignette to read that describes either a study or an analysis that has been completed or is similar to a real life example. Using the points described in the lecture, and the checklist here (<https://deon.drivendata.org/#data-science-ethics-checklist>) consider some things that could potentially be an issue. List as many as you can think of. After 10 minutes you will be asked to give a short summary to the group on what some of the potential issues are.

Vignettes

Group 1

A study of healthcare data aims to identify pneumonia risk. They consider the likelihood of someone presenting at hospital with pneumonia later dying. Their aim is to assist emergency rooms in triaging patients and identifying those who are high risk so that these patients can be admitted while the low risk patients are treated at home. They use a neural net (categorization algorithm) to predict the probability of death given baseline characteristics including demographics like age and gender, blood tests, and a medical history. The data they use is 14, 199 anonymized electronic health care records. One factor they identify is that having asthma lowers the probability of death after presenting with pneumonia. Is there an issue with using this model to assign treatment at home versus hospital treatment?

Group 2

A team of researchers at a large tech group start to develop an algorithm to identify which resumes are most similar to those that already work at the company and are successful. Their aim is to assist recruiters in identifying which job candidate would be most likely to succeed at this company. Each potential recruit is given a score which predicts how likely that candidate is to be a successful recruit. What issues surround this analysis?

Group 3

A team of researchers scrape data from a popular dating website. The dating website includes a profile picture, as well as the individual's gender identity and the gender that they are looking for in a partner. They seek to explore whether they can train a machine learning algorithm to identify individuals who are interested in the same gender as they identify as. They train an image classification algorithm to predict from profile picture whether the person is interested in same-sex relationships or not. What are the issues around this analysis?

Group 4

A US state employs a group of researchers to create an algorithm to predict the probability of reoffending given person level characteristics. They use a large dataset where baseline characteristics are known, as well as the actual rate of reoffending. They adjust for a number of important characteristics, like poverty, the number of people the individual knows in prison, age, etc. They are very careful not to adjust for protected categories like race, as it is illegal to discriminate on the basis of race in the US. What are some of the potential issues with this analysis?

Group 5

Researchers use historical crime trends in neighbourhoods (i.e., which neighbourhoods have more or less crime at different times of the year, which neighbourhoods are decreasing in crime) to fit models predicting the rate of crime in any given neighbourhood. They intend to use this model to move police into neighbourhoods that have a higher predicted rate of crime. Their intended roll out with this algorithm is to update their models regularly as new crime data comes in. What are some potential issues with this data driven policy?

- What were some of the common issues that were discussed by all groups?
- Was there consensus within the group of what some of the risks might be?
- Did you find that some of your group members thought of potential risks that you did not?

Readings

We would now like you to return to your groups. We've assigned each group either a news article or an academic article outlining at least one issue with the analysis. In the next 10 minutes, please read your article and identify the main risk the article claims is present in your vignette.

- Summarise the issue outlined in your article (some articles include potential solutions; you don't need to understand these)
- Did you identify the issue that your article did?
- Did you identify different issues than what your article did?

Return to the group and summarise.

Group 1: Caruana, Lou, Gehrke, Koch, Sturm & Elhadad (2015) Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission

Group 2 <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scrap-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G>

Group 3: Gelman, Mattson and Simpson (2018). Gaydar and the fallacy of objective measurement (<http://www.stat.columbia.edu/~gelman/research/unpublished/gaydar5.pdf>)

Group 4: Chouderchova (2017) Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments

Group 5: Ensign, Friedler, Neville, Scheidegger & Venkatasubramanian (2018). Runaway Feedback Loops in Predictive Policing

Risk Management

Almost all work with data involves some amount of risk. Often while we're learning we tend to use data that has almost no risk (such as the datasets we've been using during lectures and tutes). We'd like you to break into groups a final time to discuss how might you manage the risk that you discovered. Is there a solution for every potential risk or are some risks not worth the potential gain?