DS-UA 0202: Responsible Data Science

Credits: 4 credits

Course description and impact

The first wave of data science focused on accuracy and efficiency: on what we *can* do with data. The second wave is about responsibility: what we *should* and *should not* do. Accordingly, this technical course tackles the issues of ethics and responsibility in data science, including legal compliance, data quality, algorithmic fairness and diversity, transparency of data and algorithms, privacy, and data protection.

Data science promises to improve people's lives, accelerate scientific discovery and innovation, and bring about positive societal change. Yet, if not used responsibly -- in accordance with ethical and moral norms, and legal and policy considerations -- this same technology can cause harm on an unprecedented scale. Algorithmic changes in search engines can sway elections and incite violence; irreproducible results can influence global economic policy; models based on biased data can legitimize and amplify racist policies in the criminal justice system; algorithmic hiring practices can silently and scalability violate equal opportunity laws, exposing companies to lawsuits and reinforcing the feedback loops that lead to lack of diversity, which is both socially undesirable and can negatively impact performance of organizations. These strategic issues become more important as the economy globalizes. Therefore, as we develop and deploy data science methods, we are compelled to think about the effects these methods have on individuals, population groups, and on society at large.

The European Union recently enacted the General Data Protection Regulation (GDPR) that mandates legal protections of data subjects on the part of government entities and companies that employ algorithms and data to make decisions. The US is following suite with a plethora of local efforts, including a recently passed algorithmic transparency law in New York City that applies to City agencies. These legal framework, and many others that will soon follow, compel us to develop skills and acquire methodologies for operationalizing responsibility.

An important feature of this course is its holistic treatment of the data science lifecycle, beginning with data discovery and acquisition, through data cleaning, integration, querying, analysis, and result interpretation. For example, when considering fairness and diversity in hiring and college admissions, we will be looking at how the data was collected, cleaned and otherwise pre-processed, before analyzing the fairness of a particular classification or ranking method.

Learning objectives

After successfully completing the course, students are able to

- Construct an end-to-end case study that illustrates the role of data science in society.

- Explain the ethical and/or legal constraints in the collection and sharing of data according to a framework of the student's choice.
- Articulate the differences between various interpretations of algorithmic fairness, and relate these interpretations to the points of view of different stakeholders.
- Be conversant in the implications of current algorithmic approaches as they relate to culture and society.

Office hours

The course instructor is available two hours per week for one on one meetings with students. ***[Specifics added here for finalized syllabus]***

Topics covered

The course is structured into a sequence of lectures and accompanying assignments.

The assignment consists of labs and homeworks.

- Labs are short exercises done in class and submitted in class.
- Homeworks are longer exercises designed to take one to two weeks.

This course uses the python programming language. All labs and homework assignments are expected to be formulated as Jupyter notebooks.

Two of the homework assignments are stand-alone essay-type assignments. Additionally, writing is embedded in the labs and homeworks. During the labs, students write computer programs that analyze provided datasets. The computer programs are complex to write, as often each part must be cohesive with the others. During the homeworks, students write computer programs and provide short essays on the interpretation of the data and the implications for the data around some decision or problem that the data inform. The writing length would typically be one to two to pages for these essays.

This course does not have a required textbook. Each topic will be accompanied by required reading, as listed in the weekly schedule. In some cases, expert-level technical research papers are listed as assigned reading. These papers will be summarized as lecture notes at a level that is accessible to students.

Background Reading (required)

- Barocas & Selbst (2016) "Big Data's Disparate Impact" link (62 pages, read in weeks 1, 2, 3 of the course)
- White House Report on Big Data (2014) "Big Data: Seizing Opportunities, Preserving

Values" link (85 pages, read in weeks 4, 5, 6 of the course)

- Brauneis & Goodman (2017) "Algorithmic Transparency for the Smart City" link (74 pages, read in weeks 10, 11, 12 of the course)
- Kroll et al. (2017) "Accountable Algorithms" link (74 pages, read in weeks 13, 14 of the course)

Background Reading (optional)

- Cathy O'Neil "Weapons of Math Destruction" (book)
- Frank Pasquale "The Black Box Society" (book)
- Virginia Eubanks "Automating Inequality" (book)

A weekly schedule is below. Each week has two 75-minute lectures and one 50-minute lab. Some weeks have a homework that is due the following week.

1. Background on responsibility
   a. Topics: Introduction and overview, highlighting aspects of responsibility in data science through recent examples.
   b. Reading
      i. Friedman & Nissenbaum (1996) "Bias in Computer Systems" ACM DL link (18 pages)
      ii. Hartnett (2016) "How to Force Our Machines to Play Fair: An Interview with Cynthia Dwork" link (not paginated, about 3 pages)
      iii. Abiteboul & Stoyanovich (2016) "The Data, Responsibly Manifesto" link (not paginated, about 4 pages)
   c. Lab: refresher on Jupyter notebooks, pandas, basic data visualization.

In this week's lab, students work with supplied Jupyter notebooks that implement basic data analysis and visualization tasks. Students execute the provided programs and customize them. There is no homework assignment for this week.

2. The data science lifecycle. Data profiling.
   a. Topics: Overview of the data science lifecycle. Data profiling and validation. Is my dataset biased? Documenting data transformations. Normalization and standardization.
   b. Reading
      i. Abedjan, Golab, Naumann (2017) "Data Profiling: A Tutorial" ACM DL link (5 pages)
      ii. Stoyanovich et al (2017) "Fides: Towards a Platform for Responsible Data Science" ACM DL link (6 pages)
   c. Lab: exercises on data profiling, validation, transformation.
   d. Homework 1: implementing and documenting data profiling, validation, transformation.

In this week's lab and homework, students are introduced to interacting with SQL, and use provided python libraries to profile and transform datasets. The homework deliverable is a python Jupyter notebook that implements and documents the required transformations, and an accompanying 1-2 page written report that discusses the insights.

3. Data sharing. Anonymity and privacy.
    a. Topics: Overview of responsible data sharing. Anonymization techniques; the limits of anonymization. Harms beyond re-identification.
    b. Reading
        i. Sweeney (1997) "Weaving Technology and Policy together to Maintain Confidentiality" link (13 pages)
        ii. Foulks (1989) "Misalliances in the Barrow Alcohol Study" link (11 pages)
        iii. Ethics frameworks: The Belmont Report link (10 pages), Community Principles of Ethical Data Sharing (CPEDS) link (not paginated, about 5 pages)
        iv. Legal frameworks: General Data Protection Regulation (GDPR) link (not paginated, use as a reference)
    c. Lab: Looking at the Barrow, Alaska Alcohol Study through the lens of CPEDS.
    d. Homework 2: Written report on assigned reading (10 pages)

This week's lab and homework continue the in-class discussion of the privacy and anonymity considerations in data sharing, and of the potential harms to individuals and populations. The homework deliverable is a 10-page report that considers a recent controversial case of data sharing and use, and compares how this case would be treated by the ethical frameworks covered in class.

4. Anonymity and privacy continued.
    a. Topics: Differential privacy; privacy-preserving synthetic data generation; exploring the privacy / utility trade-off.
    b. Reading
        i. Dwork & Roth (2014) "The Algorithmic Foundations of Differential Privacy" (Chapters 1 - 6 pages, 2 - 16 pages, 3 - 37 pages) link
    c. Lab: Introducing the Data Synthesizer tool.
    d. Homework 3: Generating privacy-preserving synthetic datasets with the Data Synthesizer.

This week's lab and homework focus on privacy-preserving synthetic data generation. During the lab, students are introduced to the Data Synthesizer, an open-source toolkit, implemented in Python. The homework deliverable is an implementation of a multi-step privacy-preserving synthetic data generation scenario, subject to the data owner's requirements, that uses the Data Synthesizer library. Students submit a Jupyter notebook with their implementation, and a 1-2 page written report that discusses how their implementation meets the requirements.

5. Data cleaning.
    a. Topics: Qualitative and quantitative error detection. Missing attribute values and imputation. Outlier detection; duplicate detection. Documenting data cleaning transformations.

b. Reading
    i. Chu et al. (ACM SIGMOD 2016) "Data Cleaning: Overview and Emerging Challenges" ACM DL link (6 pages)
c. Lab: Data cleaning, profiling for bias before and after cleaning.
d. Homework 4: Data cleaning, profiling for bias before and after cleaning.

This week's lab and homework focus on detecting and mitigating data errors, and on documenting the data cleaning transformations. During the lab, students are introduced to a python library that implements common error detection and data cleaning steps. The homework deliverable is a python implementation of an error detection and data cleaning pipeline. Students submit a Jupyter notebook with their implementation, and a 1-2 page written report that discusses their implementation.

6. Midterm.
    a. Topics: Midterm review; in-class midterm exam.
    b. Reading: Lecture notes.
    c. Lab: Active study time for the midterm exam.

7. Algorithmic fairness.
    a. Topics: A taxonomy of fairness definitions; individual and. group fairness. The importance of a socio-technical perspective: stakeholders and trade-offs.
    b. Reading:
        i. Angwin, Larson, Mattu, Kirchner (2016) "Machine Bias" link (not paginated, about 6 pages)
        ii. Dwork, Hardt, Pitassi, Reingold, Zemel (2011) "Fairness through awareness" ACM DL link (13 pages)
    c. Lab: ProPublica COMPAS investigation.
    d. Homework 5: ProPublica COMPAS investigation - articulating points of view of different stakeholders.

This week's lab and homework focus on different notions of fairness, and on the trade-offs they introduce from the points of view of the stakeholders. During the lab, students are introduced to a library that implements different fairness notions (https://github.com/algofairness/fairness-comparison/tree/master/fairness or similar). The homework deliverable is a python implementation of a classifier, under alternative fairness constraints. Students submit a Jupyter notebook with their implementation, and a 1-2 page written report that discusses their implementation and the fairness trade-offs.

8. Algorithmic fairness continued.
    a. Topics: Impossibility results; causal definitions; fairness beyond classification.
    b. Reading:
        i. Friedler, Scheidegger, Venkatasubramanian (2016) "On the (im)possibility of fairness" link (16 pages)
        ii. Chouldechova (2016) "Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments" link (6 pages)
    c. Lab: Fairness in ranking.

This week's lab continues the discussion on fairness, this time for ranking tasks. Students use a fairness in ranking library, execute and modify the provided Jupyter notebook, and discuss the differences in fairness definitions for ranking vs. for classification (covered in Week 7). There is no homework assignment this week.

9. Diversity.
    a. Topics: Background on diversity in information retrieval, recommender systems, crowdsourcing; diversity models and algorithms; diversity vs. fairness; trade-offs between diversity and utility.
    b. Reading:
        i. Robinson & Dechant (1997) "Building a Business Case for Diversity" link (11 pages)
        ii. Crawford (2016) "Artificial Intelligence's White Guy Problem" link (not paginated, about 2 pages)
        iii. Drosou, Jagadish, Pitoura, Stoyanovich (2017) "Diversity in Big Data: A Review" link (12 pages)
    c. Lab: Set selection with fairness and diversity constraints.
    d. Homework 6: Set selection with fairness and diversity constraints.

This week's lab and homework build on the fairness discussion from weeks 7 and 8, and discuss diversity. During the lab, students are introduced to a python library that implements different diversity notions from Drosou et al.. The homework deliverable is a python implementation of a set selection method, under diversity constraints. Students submit a Jupyter notebook with their implementation, and a 1-2 page written report that discusses their implementation and the trade-offs between diversity and utility.

10. Transparency.
    a. Topics: Auditing black-box models; explainable machine learning; software testing.
    b. Reading:
        i. Datta, Sen, Zick (2016) "Algorithmic Transparency via Quantitative Input Influence" link (20 pages)
        ii. Ribeiro, Singh, Guestrin (2016) "Why Should I Trust You? Explaining the Predictions of Any Classifier" link (10 pages)
        iii. Galhotra, Bruin, Meliou (2017) "Fairness Testing: Testing Software for Discrimination" ACM DL link (13 pages)
    c. Lab: Generating Local Interpretable Model-Agnostic Explanations (LIME)

This week's lab starts a discussion on auditing of black-box models. Students work with a provided Jupyter notebook that implements LIME, applying this algorithm to different datasets and classification tasks, and discussing the results. There is no homework assignment this week.

11. Transparency continued.
    a. Topics: Online price discrimination, transparency in online ad delivery.
    b. Reading:
        i. Valentio-De Vries, Singer-Vine, Soltani (2012) "Websites Vary Prices,

Deals, Based on User's Information" link (not paginated, about 4 pages)
      ii. Sweeney (2013) "Discrimination in Online Ad Delivery" link (36 pages)
      iii. Datta, Tschantz, Datta (2015) "Automated Experiments on Ad Privacy Settings: A Tale of Opacity, Choice and Discrimination" link (21 pages)
   c. Lab: Executing a transparency report with the QII framework, interpreting results.
   d. Homework 7: Customizing transparency reports.

This week's lab and homework focus on producing transparency reports using a causality framework such as QII. During the lab, students execute a prepared Jupyter notebook on a given dataset, to produce a report of the influence of individual features on a classification outcome. The homework deliverable is a python implementation of a transparency report that explains the joint and marginal influence of the specified features on the classification outcome.

Students submit a Jupyter notebook with their implementation, and a 1-2 page written report that discusses their implementation.

12. Interpretability.
   a. Topics: Transparency and accountability. Legal frameworks: GDPR and the right to explanation; NYC ADS transparency law. From auditing to interpretability.
   b. Reading:
      i. Gladwell (2011) "The Order of Things: What College Rankings Really Tell Us" link (11 pages)
      ii. Yang, Stoyanovich, Asudeh, Howe, Jagadish, Miklau (2018) "A Nutritional Label for Rankings" ACM DL link (4 pages)
      iii. New York City Automated Decision Systems Task Force Report (12/2019, *forthcoming*, use as a reference)
   c. Lab: Interpreting ranked results. Working with the RankingFacts tool.
   d. Homework 8: Report on assigned reading (10 pages)

This week's lab and homework focus on interpretability. During the lab, students work with the RankingFacts tool to produce "nutritional labels" that explain ranked results. The homework deliverable is a 10-page report that discusses how the nutritional label metaphor can be applied in the scenario of a student's choice.

13. Designed at the discretion of the instructor
14. Designed at the discretion of the instructor
15. Final exam.
   a. Topics: Final review; in-class final exam.
   b. Reading: Lecture notes.
   c. Lab: Active study time for the final exam.

Course assessment

All assignments (labs and homeworks) must be entirely be the student's own submissions. Any sharing or copying of assignments is cheating and will result in an F in the course. A second cheating incident will, by CAS rules, result in a one-semester suspension from the College.

Students accumulate up to 100 points during the course.

- Up to 10 points for completing labs
- Up to 12 points for completing homeworks on time
- Up to 32 points for the midterm
- Up to 46 points for the final.

Grades will be determined using this scale:

Grade in Course	**Points Earned**
A	94 - 100
A-	90 - 93
B+	87 - 89
B	84 - 86
B-	80 - 83
C+	76 - 79
C	72 - 75
C-	70 - 71
D+	66 - 69
D	62 - 65
D-	60 - 61
F	Less than 60

Moses statement

Academic accommodations are available for students with disabilities. The Moses Center website is www.nyu.edu.csd. Please contact the Moses Center (212-998-4980 or mosescsd@nyu.edu) for further information. Students requesting academic accommodations are advised to contact the Moses Center as early as possible in the semester for assistance.