

Ethical Issues in Data Science

Chapter 6

Ethics in Data Science

- Data Ethics encompasses moral obligations of collecting, storing, protecting and using data which sometimes are sensitive, personal, emotional and behavioral
- Studies and evaluates the moral problems related to data (including generation, recording, curation, processing, sharing, uses) and algorithms (including AI, machine learning and robotics) and related practices.
- We learn about ethical problems that occur during the usage of data.
- The modern technologies generates and make use of huge amount of data collected from huge varieties of sources.
- Almost every human activities and behaviors are transformed and translated to the data to make products, decisions, betterment of services etc.

Ethics in Data Science (contd.)

- Advanced technologies like Machine Learning and AI has brought many innovation to life that makes human life better. Example use of ML to diagnosis a disease.
- Data Ethics is the center of concern for anyone who works/handles data like data analysts or data scientists or any IT professionals.

Ethical Concerns

Data Ethics concerns during:

1. Data Collection

- Data is collected through various technique like Survey, Web scrapping, social medias etc.
- What data needs to be collected and what is the privacy concern related with those data? Is the utmost concern.
- Data can be theft, shared or downloaded via TOS violation concerning data.
- Data from secondary sources or secondary use of data should be well taken care of.
- Because data collection can be repetitious, time-consuming, and tedious there is a temptation to underestimate its importance.
- Those responsible for collecting data must be adequately trained and motivated.
- They should employ methods that limit or eliminate the effect of bias
- They should keep records of what was done by whom and when.

Ethical Concerns (contd.)

2. Data Storage

- Third party storage - Is it safe? Is it secure? Who can access it? What is the mechanism of authentication and authorization?
- Hardware security - Accessibility to the computers or machines to access data. Accessibility and portability of storage devices. Can it be moved? Can it be accessed easily? Where is it stored?
- Data Security - What does the contract say? What is the level of privacy or severity?

Ethical Concerns (contd.)

3. Data Usage, Sharing and Reproducibility

- How the public data is used? Who can access it? Can it be used or reused for different research? Can it be reproduced?
- What is the terms of use for public data?
- Can the data can be shared? How is it shared?

Ethical Concerns (contd.)

4. Re-identification and Consent

- Accessing through google, something published for public usage.
- Permission to reuse the data.
- What is the privacy level of data?

5. Data Security

Limiting Access

- Locked Paper Records Offices
- Limiting access to Paper or Electronic records to appropriate personnel
- Password Protection of electronic records
- Defined privileges for electronic data users
- Firewalls to prevent outside access
- Regular Backups and proper archiving

Bias and Fairness in Data

Bias

- Machine learning model should produce a fair result.
- ML model based on data of human behavior can also have biased behaviors and tendencies.
- Cognitive biases are an obstacle when trying to interpret information
 - Can easily skew results
 - They are innate tendencies

Fairness

- Model should treat people, group or community equally irrespective of caste, religion, gender, income level, education level etc.
- Free from unnecessary and undue weighting to certain groups or viewpoints.

Common Biases

In Group Favoritism and Outgroup Negativity

In Group Favoritism

- Also called ingroup love
- Tendency to give preferential treatment to the same group they belong to.
- Very likely to occur during data collection.
- Also likely to occur during data filtering or removing irrelevant data.
- Highly impacts when data diversity is needed.

Outgroup Negativity

- Also known as outgroup hate.
- Tendency to unlike the behavior, activities or people themselves who do not belong to the group they do.
- Very likely to occur during data collection
- Likely to have covered the most of negative aspects only of outgroup community

Common Biases (contd.)

Fundamental Attribution Error

- Tendency that the situational activity or behavior are attributed as intrinsic quality of someone's character.
- These are the judged or observed pattern and is very likely to occur during data collection.
- This feeds negative data to the machine learning model resulting in biased conclusion.

Negativity Bias

- Tendency of emphasizing negative experiences over positives ones.
- This is very likely to occur during decision making.
- The negative thought about society may expect the negative conclusion from the data science projects.

Common Biases (contd.)

Stereotyping

- This is the tendency of expect a certain characteristics or behaviors without having actual information.
- This is the expectation set prior to the exploration.
- This is likely to occur during data wrangling and exploratory data analysis.

Bandwagon Effect

- Tendency to follow others because
 - Some other top ranked researcher or people did.
 - All people are doing I.e. following the mass.
- Likely to occur during data collection like same sort of data is collected based on pre-collected data or research.
- Some might expect the same result as others has inferred.

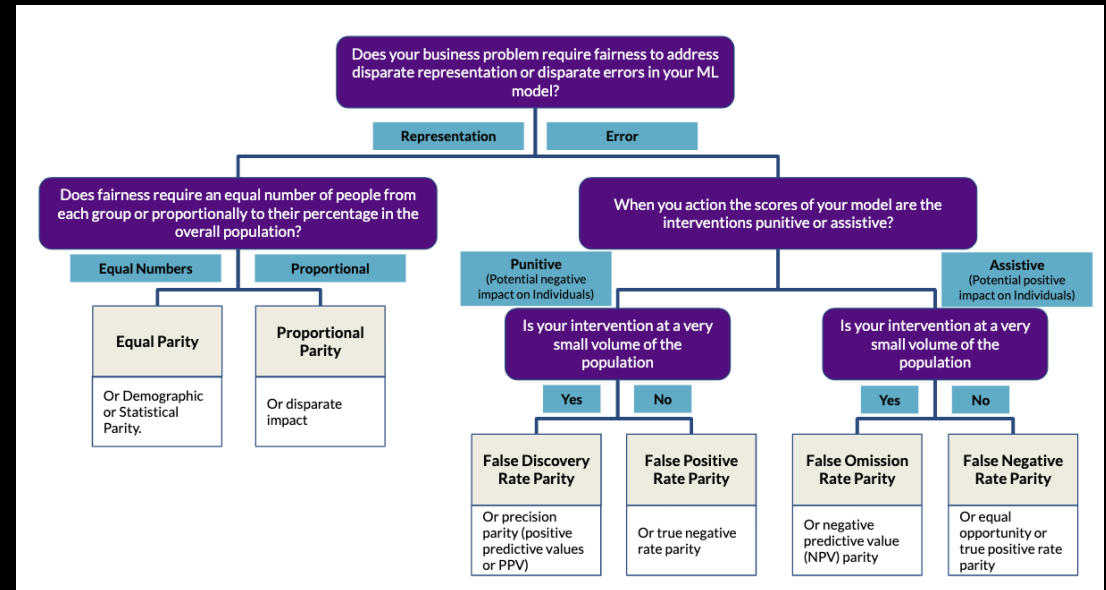
Common Biases (contd.)

Bias Blind Spot

- Our tendency not to see own personal biases.
- Likely to ignore or remain unnoticed where there are personal blind spot biases.
- Likely to occur from data collection to result analysis of data science process

Addressing Bias

- Addressing bias in data science is an extremely complex topic and most importantly there are no universal solutions or silver bullets.
- Before any data scientist can work on the mitigation of biases we need to define fairness in the context of our business problem by consulting the following:
- *As an example, imagine you want to design some ML system to process mortgage loan applications and only a small fraction of applications are by women.*



Addressing Biases (contd.)

1. Group unaware selection

- It's a preventive measure
- This is the process of preventing the bias by eliminating the factor that is likely to cause.
- For example, avoid the collection of gender to avoid bias by gender.

2. Adjusted group threshold

- Adjust any biased and unbalanced data
- Because historic biases make women appear less loan-worthy than men, e.g. work history and childcare responsibilities, we use different approval thresholds by group.

Addressing Biases (contd.)

3. Demographic Parity

- The output of the machine learning model should not depend on the sensitive demographic attribute like gender, race, ethnicity, education level etc.

Addressing Biases (contd.)

4. Equal Opportunity

- Equal opportunity fairness ensures that the proportion of people who should be selected by the model ("positives") that are correctly selected by the model is the same for each group. We refer to this proportion as the true positive rate (TPR) or sensitivity of the model.
- A doctor uses a tool to identify patients in need of extra care, who could be at risk for developing serious medical conditions. (This tool is used only to supplement the doctor's practice, as a second opinion.) It is designed to have a high TPR that is equal for each demographic group.
 - Provide equal opportunity to the diverse population.
 - Should be fair enough for the representation in sampling and treatment.
 - E.g. The representation of men and women should be same for granting loan in bank.

Addressing Biases (contd.)

5. Precision Parity

- Tune the output of model to treat the group equally.
- Male and Female should get equal salary based on the position. If machine learning model suggests lesser salary to women compared to men in same post, then such model should be tuned so that both have similar earning.
- When building a ML model, keep de-biasing in mind.

Common issues with privacy and data ethics

- Data privacy ethical issues are concerns about how data is collected, stored, shared, and used by organizations and individuals.
- Some of these issues include privacy, public safety, digital divide, working conditions, professional standard.

[Ethical Issues Related to Data Privacy and Security: Why We Must Balance Ethical and Legal Requirements in the Connected World - IEEE Digital Privacy](#)

[Ethics in Data Science and Proper Privacy and Usage of Data \(analyticsvidhya.com\)](#)

End of the Chapter