Data Ethics

Course:
INFO-6145 Data Science and Machine Learning

Revised by:
Mohammad Noorchenarboo

October 31, 2024

Contents

Current Section

What is Data Ethics?

Data ethics is a branch of ethics focusing on the evaluation of practices involving data–such as collection, creation, analysis, and sharing–that can affect individuals and society.

Objective of Data Ethics

To ensure data use respects values like transparency, fairness, and individual privacy, avoiding harm to people and society.

Example of Data Ethics in Practice

Considering the ethics of tracking a user's online activity to recommend products. Is it respectful of privacy? Does it risk making users uncomfortable?

Foundational Ethical Terms

- **Society**: A group of people organized under a system of rules to promote the well-being of members over time.
- **Morality**: The set of rules in a society about what people should or shouldn't do in various situations.
- **Ethics**: Thinking carefully and logically about what is right and wrong and examining people's behavior to understand why something may be right or wrong.

Why Ethics Matter in Data Science

Ethics in data science helps us decide not just what we can do with data but what we **should** do, ensuring that we avoid harm and act fairly.

Subjective Relativism

Subjective relativism is the idea that each person decides what is right or wrong for themselves, based on their own beliefs or feelings.

Example

One person may think it's right to collect data without asking for permission, while another person may feel it's wrong, but both feel justified in their beliefs.

Limitations of Subjective Relativism

- It blurs the line between personal preference and ethical responsibility, making it unclear when actions are truly right or wrong.
- It can fail to hold people accountable, as each person's idea of right and wrong might be very different.

Cultural Relativism

Cultural relativism is the idea that what is right or wrong depends on the beliefs and values of the culture or society in which an action occurs.

Example

In one country, it may be acceptable to collect and store data about people without asking, while in another, strict privacy rules might require asking for permission first.

Challenges of Cultural Relativism

- It's hard to apply when different cultures interact, as they may have conflicting views on ethics.
- It doesn't explain how these moral rules are created or how they change over time.

Kantianism

Kantianism is based on the idea that we should act from a sense of duty and good intentions. Immanuel Kant, who created this theory, believed that a **good will**–or the desire to do the right thing–is the only true good.

Kant's Categorical Imperative

"Act in a way that treats everyone as valuable in themselves, not just as a tool to get what we want."

Example in Data Ethics

Respecting users by seeking their consent before collecting their data shows a commitment to treating them as valuable individuals, not just as sources of data.

Kantianism

Limitations of Kantianism

Kantianism can be rigid, allowing no flexibility in rules, which can be difficult when there are conflicting duties or situations that might need exceptions.

Utilitarianism

Utilitarianism focuses on the consequences of actions, aiming to do the greatest good for the greatest number of people.

Principle of Utility

An action is right if it leads to the greatest overall happiness or benefit for the most people.

Example in Data Ethics

Sharing data on disease outbreaks with health agencies might benefit the public by helping to contain the spread of disease, even if some people's data privacy is compromised.

Utilitarianism

Challenges of Utilitarianism

- It can be hard to predict all consequences and to decide whose happiness matters most.
- Sometimes, it may justify actions that harm some people if it benefits the majority, which can be ethically troubling.

Social Contract Theory

Social contract theory is based on the idea that people collectively agree to abide by certain rules to create a society that protects individual and collective rights. These rules are typically enforced by a governing body, which is responsible for ensuring that everyone benefits from a safer and more predictable environment.

Concept of Social Contract in Data Ethics

In data ethics, social contracts involve agreeing to some limitations on data access and usage in exchange for increased security, privacy, and trust. Just as people give up some personal freedoms for the stability and safety of society, data users agree to ethical and legal standards to protect privacy and foster trust in digital systems.

Social Contract Theory

Example in Data Ethics: GDPR

The GDPR (General Data Protection Regulation) is an example of a social contract in data ethics. Under GDPR, organizations commit to transparency and fair data handling practices, like securing personal data and allowing individuals control over their information. In exchange, they gain the trust of the public, who feels more secure sharing their data within these frameworks.

Comparing Ethical Theories in Data Ethics

Each theory provides a different viewpoint on ethical decisions:

- **Subjective Relativism** - Personal views shape ethics, but lacks consistency.
- **Cultural Relativism** - Ethics depend on cultural norms, but this can create conflicts in global contexts.
- **Kantianism** - Values respect for individuals but can be inflexible.
- **Utilitarianism** - Aims to maximize benefits for the most people but can overlook individuals' rights.
- **Social Contract Theory** - Balances personal rights with societal rules, supporting a fairer society.

Choosing an Ethical Framework

Choosing a framework depends on the context and goals in data ethics. Each theory offers valuable insights for deciding how to act responsibly.