University of Lancashire

Data Science: Ethics

Karen O’Shea

Where opportunity creates success

Learning Objectives

University of Lancashire

Evaluate the principles of AI ethics and their importance in research design

Discuss ethics and apply to emerging systems

Starter Activity

- Why consider ethics?
- What do they mean to you?
- How does ethics differ from ‘morals’?
- Is there a ‘right’ or ‘wrong’ to any ethical dilemma?

Last Week: The Turing Test Approach

- Read Turing's original paper on AI Turing:1950. In the paper, he discusses several objections to his proposed enterprise and his test for intelligence.

Discussion points:

- Which objections still carry weight?

- Can you think of new objections arising from developments since he wrote the paper?

Reference: Based on: (Russell and Norvig, 2014).

Key Concerns on AI Computer Intelligence

- Impact on Jobs
- Privacy
- Bias and discrimination
- Accountability

Ethics (1)

Video: UNESCO – Why is ethics important in the development of Emerging Technologies? https://www.youtube.com/watch?v=HzYG56HLxb

Document: Preliminary study on the ethics of AI
https://unesdoc.unesco.org/ark:/48223/pf0000367823

Ethics (2)

University of Lancashire

- Data is the foundation for machine learning algorithms
- Human beings and cognitive bias
- Data security and integrity
- Impact research design
- Legal exposure and costly penalties

Ethics (3)

Innovation outpaces government regulation in new, emerging fields.

Protocols to avoid any infringements on human rights and civil liberties.

Establishing Principles.....

Principles on Ethics

- Belmont Report – guide to ethics within experimental research and algorithm design.

https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c_FINAL.pdf

Three Main Principles

- **Respect for Persons** – all individuals, variety of personal circumstances, and risk during experiment design

- **Benefit** – for example, do no harm. Intention to do good. Consider bias/political leanings, race, gender, etc.

- **Justice** – fairness and equality. The Belmont Report offers five ways to distribute burdens and benefits, which are by: Equal share; Individual need; Individual effort; Societal contribution; Merit.

Establishing AI Ethics

- **Governance**: Companies can leverage their existing organisational structure to help manage ethical AI. If a company is collecting data, it has likely already established a governance system to facilitate data standardisation and quality assurance.

- **Explainability**: Machine learning models, particularly deep learning models, are frequently called “black box models” as it’s usually unclear how a model is arriving at a given decision.
  - Explainability seeks to eliminate this ambiguity around model assembly and model outputs by generating a “human understandable explanation that expresses the rationale of the machine”.

Reference:

IBM, AI Ethics, https://www.ibm.com/cloud/learn/ai-ethics

University of Central Lancashire
UCLan

Cyber Security: Poison in the Well

Further Challenges:

- Shared resources of AI (Machine Learning):
- Open source data sets
- Pre-trained models
- Machine learning libraries

Read report here:

https://cset.georgetown.edu/publication/poison-in-the-well/

Scenario

Criminal Justice System – Recidivism

- Should AI be used to predict the likelihood of a person re-offending and if so, to what extent should the data influence the decision.

Consider:

- data;
- decision-making processes;
- impact on human factors;
- positive/negative outcomes;

For and against – some thoughts…

- Efficient in decision-making; vast amounts of data; potential to target individuals for rehabilitation.

- Bias; loss of human judgement; human factors (personal circumstances); transparency of the decision

Assessment

Learning Objectives

University of Lancashire

Evaluate the principles of AI ethics and their importance in research design

Discuss ethics and apply to emerging systems