

PS 643: Introduction To AI and Data Policy, Fall 2024:

Lecture Notes [03-Sep-2024]

Name : Akash Pal

Roll No : 23m2158

AI in Job Suitability and Surveillance:

- AI systems are often claimed to predict job suitability or employee productivity, but these claims are questionable.
- Some companies, especially in countries like South Korea, teach employees how to behave or smile to avoid suspicion from AI assessments.
- Surveillance systems often serve more to create the illusion of security rather than genuinely improving safety.

AI and Productivity:

- Fear and control are often used to drive productivity rather than the actual effectiveness of AI tools.
- The psychological impact of being monitored or assessed by AI can lead to increased productivity, even if the AI itself doesn't work as claimed.

Fraud in AI:

- Many AI systems, particularly those in the market, are driven by fear and greed, drawing parallels to the historical "snake oil salesmen."
- There is public confusion about the capabilities of AI, leading to large investments in technologies that may not work as promised.

Categories of AI:

- **Genuine Progress:** AI that works effectively (e.g., facial recognition or speech recognition) but may still have ethical issues.
- **Useful but Limited:** AI systems that solve problems but may not work in all situations, like recommendation systems.
- **Complete Failures:** AI systems that try to predict complex social behaviors or outcomes (e.g., predicting job performance or crime) but are inherently flawed due to the complexity of human behavior and society.

AI and Social Outcomes:

- Attempts to predict human behavior using AI (e.g., for predicting criminal recidivism) often fail because AI cannot account for the complexities of society.
- Case study of the COMPAS system: This AI, used in the U.S. to predict whether criminals will reoffend, was found to be biased and less accurate than simple statistical methods.

Philosophy of Ethics and AI:

- **Normative Ethics:** Focuses on determining what actions are right or wrong, with different schools of thought like virtue ethics, deontological ethics, and consequentialism.
- **Virtue Ethics:** Evaluates a person based on their moral character.
- **Deontological Ethics:** Judges actions by their adherence to rules or duties.
- **Consequentialism:** Assesses actions based on their outcomes, aiming to maximize happiness or utility (utilitarianism).

AI and Policy:

- AI's impact on policy decisions can be significant, with examples like smart cities and automated decision-making.
- There are concerns about the transfer of decision-making power from human experts (e.g., judges) to AI systems developed by private companies, raising issues of accountability and transparency.

Ethical Issues in AI:

- The misuse of AI can lead to privacy violations, especially when sensitive data is used to predict social outcomes.
- AI systems that try to make decisions about welfare distribution or job performance often perpetuate biases or make incorrect predictions.