UNIVERSITY PARTNER

UNIVERSITY OF WOLVERHAMPTON

HERALD COLLEGE KATHMANDU

Concepts and Technologies of AI
(5CS037)

Project Report
Ethics of AI

Student Id : 2226675
Student Name : Roshish Shrestha
Group : L5CG4
Module Leader : Mr. Siman Giri
Asst. Module Leader : Mr. Anmol Adhikari
Instructor : Mr. Rohit Parajuli
Cohort : 8

Submitted on : 21st January, 2023

Abstract

What kind of thought processes does a computer have? How does a computer acquire knowledge? Can computers learn in the same way as people? Will computers ever have consciousness? These are just a few of the questions that philosophers and scientists have been pondering for centuries. It's debatable whether robots, computers, and other devices have moral standing. Some claim that because these machines are incapable of having any kind of self-awareness, they will never have moral status. This essay explores the ethical implications of AI and its potential impact on society. As the field of AI continues to advance, it is increasingly likely that computers and robots will become more intelligent and autonomous in their interactions with humans raising important ethical concerns about the responsibility and accountability of AI systems. The ethical considerations of AI are becoming increasingly important and it is important that these are addressed in order to ensure that future developments in AI are aligned with human values and respect human rights.

Table of Contents

Introduction

Ethics

Ethics is the study of morality and personal values that guide individuals in determining right and wrong behavior. Every person has their own set of ethical beliefs which they use to make decisions. The study of ethics in relation to AI has been gaining popularity over recent years due largely because many people are concerned about this technology becoming too powerful for humans' control over their own lives. Hence, why there exists such a need for better ways for us humans who use AI every day here today, especially given all these new developments coming along lately.

Ethical AI

Ethical AI is the development of AI system adhering with ethical principles and values. The world is changing at a rapid pace, and we need to be able to predict and respond to those changes. AI can help us do exactly that: make better decisions and solve problems faster than humans can. The rise in the use of AI in various industries has led to an exponential increase in ethical dilemmas. Handling ethical dilemmas becomes a difficult task when we encounter quality tradeoffs between effectiveness and fairness.

Why should we build Ethical AI?

We should build Ethical AI because it ensures that the technology is used in a way that respects human rights and values. It also aids to prevent harm and discrimination caused by biased algorithms and data. Furthermore, prioritizing ethical considerations in AI development can help build trust and confidence in the technology, which is critical for its successful adoption and integration into society.

Major Ethical Dilemmas

As with any new technology, there are ethical and moral considerations that must be taken into account. In my opinion, some of the major ethical and moral debates associated with AI include:

Automated decisions / AI bias: AI systems can contain biases, just like humans, since they are created by humans. These biases can lead to unfair decisions and discrimination towards minority. This can occur when developers unknowingly create biased systems or when the training data used does not fairly represent all groups. To combat this, it is important for organizations to strive to minimize biases in their AI systems through testing and identifying potential biases in data, models, and human use of algorithms.

Autonomous things: The use of self-driving vehicles raises ethical concerns, such as liability and accountability. One example is the 2018 fatal accident involving an Uber self-driving car. Despite the fact that the accident was avoidable and that the safety driver was distracted, the company was not held criminally liable for the death. This incident begs the question of who should be held accountable in similar situations involving self-driving cars.

Unemployment and income inequality due to automation: Many people are concerned that AI will cause widespread unemployment and increased income inequality. Many believe AI will take their jobs within 5 years and experts predict that by 2030, AI could replace 30% of human jobs, potentially resulting in the displacement of 400-800 million jobs. Studies also show that automation has significantly reduced wages for workers in routine tasks by 50-70% since 1980, leading to increased income inequality.

Initiatives in Ethical AI

The field of AI is rapidly advancing and has the potential to greatly benefit society. However, despite the progress and initiatives made in this field internationally, Nepal has yet to make a notable contribution to the ethical development of AI. One of the main reasons for the lack of ethical considerations in the use of AI in Nepal is the absence of laws and guidelines. Currently, there are no legal regulations in place to govern the implementation and use of AI in Nepal. This lack of oversight allows organizations and individuals to develop and implement AI systems without considering the potential ethical consequences. Some notable initiatives in Ethical AI internationally are:

Robot Ethics Charter is a set of principles to guide the development of AI, developed with input from experts in ethical AI. It outlines several ethical issues that need to be considered when developing AI systems.

OECD Principles on AI are a set of guidelines developed by the OECD and its member countries to promote responsible development and use of AI. The principles address ethical considerations such as transparency, privacy, fairness, respect for human rights, security, communication and accessibility for people with disabilities.

Generic Framework

Based on the above international notable initiatives in Ethical AI, some generic frameworks suggested on building fair, explainable and ethical AI are:

- AI systems should not be able to make decisions independently without human supervision nor should they be able to take actions which would cause harm or death (intentionally or unintentionally).
- AI system should behave appropriately in public spaces so as not to cause discomfort among other humans.
- AI developers are ethically obliged to be transparent in a structured and accessible manner.
- AI research even if it takes place in private, for-profit companies, should to be publicly shared.
- AI developers and businesses must explain how their algorithms make their predictions to avoid ethical problems brought on by inaccurate predictions.

The OECD have also developed principles for ethical AI that cover the following areas:

- **Inclusive growth, sustainable development and well-being**: Ensuring that the benefits of AI are shared across society and do not harm the environment or negatively impact human well-being.
- **Human-centered values and fairness**: Ensuring that AI systems align with human values and respect human rights, and do not perpetuate bias or discrimination.
- **Transparency and explain-ability**: Ability for users to understand how an AI system makes decisions and for developers to be accountable for the system's actions.
- **Robustness, security, and safety**: Ensuring that AI systems are reliable and safe, and that they do not pose a threat to security or cause harm.
- **Accountability**: Responsibility and accountability of AI developers and users for the actions and consequences of the AI systems they create and use.

Discussion

In conclusion, AI has the potential to transform our world for the better, but also poses significant ethical challenges. As AI becomes more powerful and sophisticated, so too does the potential for misuse. In my opinion, now more than ever, we need to ensure that AI is developed in an ethical way and benefits everyone equally. It is important that we continue to prioritize ethical considerations in the development and implementation of AI to ensure that the technology is used for the betterment of society. By adhering to the regulations and guidelines while designing and developing AI systems, we can be sure that we are not headed (Bird, 2020) into a dystopian AI led future.

References

Dilmegani, C., 2022. *Top 9 Ethical Dilemmas of AI and How to Navigate Them in 2023.* [Online]
Available at: https://research.aimultiple.com/ai-ethics/
[Accessed 20 January 2023].

New Scientist and Afp, 2007. *South Korea creates ethical code for righteous robots.* [Online]
Available at: https://www.newscientist.com/article/dn11334-south-korea-creates-ethical-code-for-righteous-robots/
[Accessed 20 January 2023].

Wakabayashi, D., 2018. *Self-Driving Uber Car Kills Pedestrian in Arizona, Where Robots Roam.* [Online]
Available at: https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html
[Accessed 20 January 2023].

Parliament, European, Directorate-General for Parliamentary Research Services, J Fox-Skelly, E Bird, N Jenner, A Winfield, E Weitkamp, and R Larbey. 2020. *The Ethics of Artificial Intelligence: Issues and Initiatives*. European Parliament.

OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449