

UNIVERSITY PARTNER



Concepts and Technologies of AI (5CS037)

Project Report [Ethics of AI]

Student Id	: 2228160
Student Name	: Ritesh jung lama
Group	: L5CG4
Module Leader	: Siman Gir
Asst. Module Leader	: Anmol Adhakari
Instructor	: Rohit Parajuli
Cohort	: 8
Submitted on	: 21/01/2023

Abstract

AI-based technology has brought many advancements to society, but it also poses ethical and moral concerns such as data biases, data security, and data privacy. As AI continues to evolve, it's crucial to understand and address these issues. The concept of "machine ethics" has been around for some time, but there is still much work to be done in this field. To build ethical AI, principles, rules, and guidelines must be established for the AI to respond to problematic situations. There are several national and international initiatives that have been taken in order to build fair, explainable and ethical AI, however, it is still a developing field and require continuous research, development, and monitoring to address the ethical and moral issues that arise with the integration of AI in our society.

Contents

Introduction	3
Review of Major Ethical Dilemmas	4
Initiatives in Ethical AI	5
Describe Generic Framework	6
Conclusion	7

Introduction

To sum up, Knowing ethics comes from studying morals. It is considered with what is good and what is bad (i.e right or wrong). A set of moral guidelines is the simplest definition of ethics. They have an impact on the decisions people make and how they spend their lives.

Ethics are reliant on a wide range of factors and are subject to modest changes in any one of them. Religious convictions, social mores, societal standards, feelings, and other considerations are only a few of them.

About Ethics of AI, according to different opinions of people they feel that humans are still not ready for AI. Ethical AI refers to the use of artificial intelligence that adheres to clear ethical principles, such as protecting individual rights, maintaining privacy, avoiding discrimination, and preventing manipulation. In order to distinguish between appropriate and inappropriate uses of AI, ethical AI places a strong emphasis on ethical considerations. Organizations that implement ethical AI have clearly defined rules and review processes in place to ensure compliance with these standards. The wrong use of AI can be harmful for humans. So, we must build AI which can take decisions on what is right and what is wrong.

Review of Major Ethical Dilemmas

A difficult situation known as an ethical dilemma frequently entails a mental struggle between two morally and practically necessary things.

One of the major ethical issues with AI is the potential for bias to be introduced during the training process. This can happen when data used to train the AI is itself biased, and can result in the AI also exhibiting bias. This is a concern because, while AI is created with the intent to be fair and just, it is ultimately designed and programmed by humans, who may unintentionally introduce bias based on factors such as race, gender, or religion. This can lead to flawed AI functionality.

As an experiment, consider training an AI system to recognize faces using a dataset where the majority of the faces belong to a single race or gender. Due to a lack of sufficiently diverse training data, this can cause the AI system to perform poorly or erroneously when recognizing faces of different races or genders. This is an illustration of bias being introduced into the AI training process by the data used, and it may cause the AI system to act biasedly.

Initiatives in Ethical AI

To ensure that AI is fair, simple to comprehend, and safe to use, numerous groups and organizations are working.

The National Artificial Intelligence Research and Development Strategic Plan is a strategy used in the US to ensure that AI is reliable and responsible. The Ethics Guidelines for Trustworthy AI are a set of principles that are utilized in Europe to ensure that artificial intelligence is created and applied in a fair and reasonable manner.

In order to ensure that AI is created and utilized in a way that is safe for people, the Organisation for Economic Co-operation and Development (OECD) has also created a set of guidelines known as the AI Principles. The Global Partnership on AI, an organization created by the UN, strives to ensure that AI is utilized in a way that benefits everyone.

In order to ensure that AI is utilized fairly and in a way that benefits everyone, these groups and organizations are attempting to prevent it from hurting individuals or society as a whole. Additionally, they strive to build AI's credibility and make it simple to understand.

Describe Generic Framework

In order to ensure that AI is fair, understandable, and safe to use, numerous groups and organizations are working to achieve these goals. For how to achieve that, they are developing frameworks and rules of thumb.

The National Artificial Intelligence Research and Development Strategic Plan, for instance, aims to ensure that AI is trustworthy and accountable. The Ethics Guidelines for Trustworthy AI are another European initiative that works to ensure that AI is created and applied in a fair and reasonable manner.

To make sure that AI is created and used in a way that is safe for people, the Organisation for Economic Co-operation and Development (OECD) has created a set of guidelines called the AI Principles. The UN has also set up a group called the Global Partnership on AI to make sure that AI is used in a way that is good for everyone. These organizations are working to make sure that AI is fair and beneficial to all, without any negative impacts on people or society. They also want to make sure that people can trust and understand AI.

Conclusion

Making sure that AI is developed and used ethically is crucial, in my opinion. This emphasizes that it should be in line with our moral principles and beneficial to society as

a whole. There is a possibility that the technology could be applied in ways that are prejudiced or harmful if ethical AI is absent.

I also think that ethical AI development needs to be a team effort. It should be handled by a number of different stakeholders, including business, government, academia, and civil society, who collaborate to exchange best practices and develop research, rather than by just one group or institution.

I find it incredibly exciting that there are programs aiming toward ethical AI, such as the National Artificial Intelligence Research and Development Strategic Plan in the US and the Ethics Guidelines for Trustworthy AI in the EU. Additionally, institutions like the United Nations and the Organisation for Economic Co-operation and Development (OECD) have their own standards and frameworks for ethical AI.

In conclusion, I believe it is essential to ensure that AI is being created and deployed in an ethical manner, and that this involves cooperation from numerous parties.