Herald College, Kathmandu

Concepts and Technologies of AI

5CS037

Assignment -2.
Write a Comprehensive Report on: "AI: Balancing Innovation with Ethical Integrity: Opportunities and Challenges across various fields."

December 17, 2024

Contents

1 Assignment Details and Submission Guidelines

1.1 Assignment Details:

Due	Marks	Submission
Jan - 3rd.	10	A report with $1200 \pm 10\%$ words. for details see page no. 3 and 4

1.2 Plagiarism and AI Generated Content

Plagiarism of more than 20% and any AI-generated content found in the report will be reported for academic misconduct. Thus, we highly encourage you to submit your original work.

1.3 Submission Guidelines:

- This assignment must be completed individually.
- What to Submit?
  - You are expected to submit a PDF file with valid plagiarism score named after following naming convention.
  - Please print a Plagiarism report generated from Canvas and ask your respected instructor to sign it for you. {Plagiarism allowed: 20%}.
- Where to submit?
  - Designated portal opened on Canvas or as instructed by your instructor.
- **Naming Conventions:** You are supposed to strictly follow the naming conventions, and any file that does not follow the naming conventions will be marked as "0".
  - Example: `00000_ABC Sharma.pdf` {For Report.}

The Final Date for submission is: 3rd **Jan.** **No Late submissions are not allowed.**

2 About Assignment:

In this assignment, you are expected to write a comprehensive report on "AI: Balancing Innovation with Ethical Integrity: Opportunities and Challenges across various Fields". In the report you are expected to explore the ethical and moral constraints posed by the transformative power of modern AI-powered technologies and applications across various industries and fields.

You are required to critically analyze and rigorously examine the challenges presented by these technologies, particularly with regard to issues such as bias, privacy, accountability, societal impacts, and more. Emphasis should be placed on proposing strategies for fostering responsible and ethical development, ensuring that AI serves as a catalyst for equitable and sustainable progress across diverse domains.

2.1 Learning Outcomes:

Following are some of the learning outcomes but are not limited to:

1. Understanding Ethical and Moral Issues in AI.

   - Identify and critically analyze the perceived ethical and moral challenges associated with AI technologies, such as bias, privacy violations, lack of accountability, and societal impacts.

2. Principles and Guidelines for Ethical AI.

   - Explore general and common ethical principles, guidelines, and regulatory frameworks designed to address and mitigate ethical and moral concerns in AI development and application.

3. Characteristics of Ethical AI.

   - Examine the essential features and characteristics that define an ethical AI, including fairness, transparency, accountability, and human-centric design.

4. Building and Adhering to Ethical AI Practices.

   - Learn strategies and methodologies for adhering to ethical principles in AI development to ensure the creation of responsible, equitable, and sustainable AI systems.

3   Division of the Task.

Your Report must have at least following sections and description:

1. Abstract [100 - 150 Words.]

An abstract must be fully self-contained and make sense by itself, without further reference to outside sources or to the actual paper. It highlights key content areas, your research purpose, the relevance or importance of your work, and the main outcomes.

2. Introduction [250 - 300 Words.]

In this section you could briefly explore and examine answer to following questions:

1. What are the perceived ethical and moral issues with AI?
2. What are the general and common ethical principles, guidelines and regulations that can attenuate these ethical and moral issues with AI?
3. What are some of the necessary features and characteristics of an ethical AI?
4. How to adhere to the ethics of AI to build ethical AI?

3. Review: Major Ethical Dilemmas and Moral Questions in AI [550 - 650 Words.]

This is the major section of your report. You will be randomly assigned one of the following topics by your instructor. Explore and critically analyze the topic you are assigned. Topics are:

1. **AI Discrimination, Decision Making, and Privacy Ethics:**
   - Explore the impact of bias in AI algorithms and ethical dilemmas in decision-making across sectors like healthcare, finance, and criminal justice.
   - Discuss the need for transparent, accountable AI systems and the responsibility of developers and organizations for their outcomes.
   - Investigate the ethical challenges of preserving privacy in AI-driven processes.

2. **Ethical Implications of AI in Creativity and Employment:**
   - Explore ethical issues surrounding AI-generated content in fields like art, music, and literature, focusing on copyright, attribution, and ownership.
   - Examine the impact of AI on employment, discussing potential solutions to mitigate job displacement and promote workforce adaptation.
   - Examine how AI's impact on employment intersects with creativity, discussing potential solutions to job displacement while promoting AI-human collaboration in creative industries.
   - Investigate the role of AI in augmenting human creativity and its ethical implications for workforce adaptation and skill development.

3. `AI For Social Good:`
   - Analyze the environmental impact of training large AI models, focusing on energy consumption and the ethical considerations related to sustainability.
   - Explore the need for international cooperation and governance frameworks to address global ethical concerns in AI, discussing challenges and potential solutions for equitable regulation.
   - Investigate how AI can either contribute to or exacerbate existing social inequalities, and explore strategies to ensure that AI systems promote social justice and inclusivity.
   - Investigate the role of AI in addressing global challenges such as climate change, poverty, and access to education, highlighting both the potential benefits and ethical concerns.

4. `LLMs and Eduction:`
   - Examine the ethical implications of bias in LLM-generated educational content, ensuring that AI tools promote inclusivity and diversity in learning materials.
   - Analyze the role of LLMs in automating teacher-student interactions, such as tutoring systems, while considering the ethical risks of reducing human teacher involvement and the quality of feedback.
   - Explore the ethical challenges of using LLMs in assessment and grading, focusing on fairness, accuracy, and transparency in automated evaluations, and the potential for AI to influence educational outcomes.

5. `AI and Modern Warfare:`
   - Examine the ethical concerns surrounding the use of AI in autonomous weapons systems, focusing on accountability in decisions made by AI-driven machines in conflict zones and the potential for unintended harm.
   - Explore the implications of AI-enabled surveillance and intelligence gathering, discussing the balance between national security and individual privacy, and the ethical responsibility of nations in the use of such technologies.
   - Analyze the potential for AI to escalate conflict, including the ethical dilemma of autonomous weapons potentially acting without human intervention, and the danger of miscalculation in military engagements.

For any of the topics assigned to you, in addition to the questions above, also explore national and international initiatives aimed at building ethical AI. Suggest a generic framework that could be applied to address any of the topics provided. Please try to be original and give your own understandings.

4. Discussion [150 - 250 Words.]

Based on your research, and findings provide your view on importance of building ethical AI.

4 After Submission.

Once you made your final Submission, Please save and download the Plagiarism report and request your Instructor to sign it.

**** GOOD LUCK *****