

AAI301 Student Study Guide

Seminar 1: Introduction to Generative AI and Ethical Foundations

Overview

This guide is designed to help you understand and apply the core concepts covered in **AAI301 Seminar 1**. If you missed the seminar or need a refresher, this document provides the essential theoretical foundations and practical activities to ensure you are up to speed.

Learning Objectives

By the end of this session, you should be able to:

- Define Generative AI and differentiate it from other forms of Artificial Intelligence (Discriminative AI).
- Identify key ethical frameworks relevant to AI technology.
- Recognize why ethical principles are critical in the development and deployment of generative AI.

Key Concepts

1. Artificial Intelligence (AI) vs. Generative AI (Gen AI)

- **Traditional AI (Discriminative AI):** Focuses on analyzing existing data to make predictions, classifications, or decisions.
 - *Analogy:* Think of it as an **Art Critic**—it looks at art and tells you if it is a Van Gogh or a Picasso.
 - *Examples:* Spam filters, Netflix recommendations, FaceID.
- **Generative AI:** A subset of AI focused on creating new content (text, images, music, code) that did not exist before.
 - *Analogy:* Think of it as an **Artist**—it studies existing art to create a brand-new painting.
 - *Examples:* ChatGPT, DALL·E, GitHub Copilot.

2. How Gen AI Works

Gen AI uses **Neural Networks** (like Transformers, GANs, and Diffusion Models) to learn patterns from massive datasets.

- **Training:** The model ingests vast amounts of data (books, images, code).
- **Pattern Recognition:** It learns the underlying structure, style, and relationships (e.g., understanding that "cat" often appears near words like "meow" or images of whiskers).
- **Generation:** It predicts the next piece of information to create new outputs that plausibly belong to the original dataset.

3. AI Ethics: Law vs. Ethics

One of the most critical takeaways is the distinction between what is legal and what is ethical.

- **Law:** "What you *can* or *cannot* do." (Enforced by penalties).
- **Ethics:** "What you *should* do." (Enforced by values and social norms).
- **Example:** It might be legal to train an AI on public social media photos, but it may be unethical due to lack of consent and privacy violations.

4. The "Rose" Case Study (Ethical Dilemma)

In class, we analyzed the generation of a rose image to understand ethical pitfalls:

- **Bias:** If an AI is trained mostly on red roses, it may fail to generate yellow or white roses, reinforcing stereotypes.
- **Copyright:** If the AI generates an image that looks exactly like a copyrighted painting, is it theft?
- **Hallucination:** The AI might generate a biologically impossible flower (e.g., a rose-sunflower hybrid) presented as real, leading to misinformation.

5. Core Ethical Principles

We established six pillars for responsible AI:

1. **Fairness & Non-discrimination:** Avoiding bias against specific groups.
2. **Responsibility & Accountability:** Determining who is answerable for AI errors.
3. **Transparency & Explainability:** Understanding *how* the AI made a decision.
4. **Privacy & Data Protection:** Safeguarding user data.
5. **Nonmaleficence & Beneficence:** "Do no harm" and "Do good."
6. **Human Autonomy:** Ensuring humans remain in control.

Practical Exercises (Catch-Up Activities)

In class, we explored several tools. Please try these on your own to familiarize yourself with the capabilities of Gen AI.

Exercise 1: Image Classification (Discriminative AI)

- **Tool:** [Teachable Machine by Google](#)
- **Task:** Train a simple model to distinguish between two objects (e.g., "Thumbs Up" vs. "Thumbs Down" or "Pen" vs. "Phone") using your webcam.
- **Goal:** Understand how traditional AI requires labeled data to classify inputs.

Exercise 2: Text Generation & Tone Shifting

- **Tool:** [Google Gemini](#)
- **Task:** Find a serious news article (e.g., about climate change). Use the following prompts:
 1. "Write a *hilarious* 100-word essay based on this article: [Insert Link]"
 2. "Write a *formal* and *alarmist* 100-word essay based on this article: [Insert Link]"
- **Reflection:** Observe how changing the tone alters the perception of the story.

Exercise 3: Image Generation (Prompt Engineering)

- **Tool:** [Leonardo.ai](#)
- **Task:** Use the following prompt:
 - "Generate an image, in watercolor style, capturing the atmosphere of a soccer match in a crowded stadium. Showcase the energy and excitement, as two teams engage in the competition. The scene is alive with action, seen from the top stand of the stadium."
- **Goal:** Experiment with adding adjectives (e.g., "cyberpunk style," "gloomy weather") to see how the output changes.

Exercise 4: Voice Cloning

- **Tool:** [ElevenLabs](#)
- **Task:** Record your voice reading a short poem. Use the "Voice Changer" feature to transform your voice into a different character.
- **Reflection:** How does it feel to hear your words spoken by a voice that isn't yours?

Exercise 5: Code Generation

- **Tool:** [Claude AI](#)
- **Task:** Enter the prompt: "Create a snake game that can be played here."
- **Goal:** Witness how Gen AI can generate functional, deployable code in seconds.

Preparation for Next Class (Seminar 2)

In Seminar 2, we will dive deeper into the specific impacts of AI on society.

- **Topic:** Bias, Fairness, Responsibility, and Accountability.
- **Preview:** We will move from *identifying* ethical principles to *analyzing* real-world failures, such as:
 - **Bias:** How AI in policing or hiring can discriminate against minorities.
 - **Fairness:** The difference between "Equal Opportunity" and "Demographic Parity."
 - **Accountability:** Who is to blame when a chatbot gives bad advice? The developer, the user, or the school?

Resources & Tool Access

- **Text:** [ChatGPT](#) | [Gemini](#) | [Claude](#)
- **Images:** [Leonardo.ai](#) | [Runway](#)
- **Audio:** [ElevenLabs](#)
- **Classification:** [Teachable Machine](#)

Self-Assessment Questions

1. What is the primary difference between discriminative AI and generative AI?
2. Why might it be legal but unethical to use certain data to train an AI model?
3. List three of the six core ethical principles discussed in the seminar.

4. In the context of the "Rose" example, how does training data bias affect the output?

Need Help?

If you have trouble accessing the tools or understanding the ethical frameworks:

- **Email:** Rudy005@suss.edu.sg
- **Office Hours:** [Week days / Between 12pm to 1pm]
- **Discussion Forum:** [to be shared later]