Task 1: Understanding & Exploring Data Bias

**Step 1: Ask ChatGPT to generate an example of a left-handed person writing.**

**Step 2: Identify if there’s bias in the response (e.g., if it assumes most people are right-handed).**

**Step 3: Research other common AI biases and document them.**

**📌 Expected Output: A list of biases found in AI-generated responses.**

AI models, including Large Language Models (LLMs), can exhibit **various biases** based on the data they are trained on.

## Left-Handed Writing Bias

#Understanding the Bias

AI models, including ChatGPT, are trained on large datasets that often reflect historical and societal norms. One such bias is the assumption that right-handedness is the default.  
If we ask an AI to describe a left-handed person writing, the response might:  
- Assume that left-handedness is uncommon or unusual.  
- Emphasize challenges(e.g., smudging ink, struggling with certain tools).  
- Use words like 'difficulty’,'awkward’,or 'adjustment’,which frame left-handedness as a problem rather than a normal variation.  
This bias originates from historical trends, where right-handed writing was promoted in education, and left-handed individuals were often forced to switch.

Below are some common types of biases found in AI-generated responses:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | **Bias Type** |  |  | | --- | |  | | **Description**   |  | | --- | |  | | **Example**   |  | | --- | |  |      |  | | --- | |  | |
| Selection Bias | AI training data may overrepresent some groups while underrepresenting others. | Facial recognition works better on lighter skin tones due to imbalanced datasets. |
| Gender Bias | AI may reinforce gender stereotypes. | A nurse" is often assumed to be female, while "a doctor" is assumed to be male. |
| Cultural Bias | AI models are often trained on Western-centric data | AI might assume Christmas is universally celebrated but ignore other cultural holidays. |
| Automation Bias | Users may trust AI responses without verification. | AI-generated medical advice may be taken as absolute truth, even when incorrect. |
| Historical Bias | AI inherits biases present in historical data. | Job hiring AI might favor men for tech roles because past hiring data was male-dominated. |

**How to Reduce Bias in AI Models?**

* **Diversify Training Data:** Include a wide range of sources and perspectives.
* **Bias Detection Tools:** Use AI fairness frameworks (e.g., IBM AI Fairness 360, Fair learn).
* **Human Oversight:** Regularly review and update AI responses.
* **Transparency:** Clearly state AI limitations and potential biases.