Exercise: Analyze the Risks of Using a Foundation Model

Lesson Downloads

In today's digital age, the use of foundation models has become increasingly prevalent in various industries. However, it is crucial to analyze the risks associated with these models to ensure their ethical and responsible implementation. This analysis involves examining potential biases that may be embedded within the foundation models, the privacy concerns that arise from their usage, and the impact on labor and the environment. In this exercise, you will dive deeper into two topics of your choice, one suggested by a foundation model and another you come up with yourself.

Ask an LLM

LLMs these days will gladly offer lists of risks that need to be kept in mind while using them. It is a relatively new technology, after all, and the creators of these LLMs want to get in front of any problems before they occur. Will they be successful? That's for you and the future to decide.

Let's write a prompt

Begin by writing a prompt (or prompts) that we will ask the LLM about the risks of using it. The prompt does not have to be long but it can be. You might reflect on what things are important to you and use this to help you craft your prompt.

Your reflection

Prompt: "What are the potential risks of using large language models (LLMs) in various domains, including ethics, privacy, misinformation, security, bias, and overreliance? Discuss real-world examples of these risks and how they have impacted individuals, businesses, or society. Additionally, suggest mitigation strategies to address these risks while leveraging the benefits of LLMs."

Things to think about

Great! Now let's ask the LLM your question and review what it says,

Choose one risk suggested by the LLM

Did the LLM suggest multiple risks? If not, try to reword your prompt so that it will return a list of risks. Then, when it does, choose one of the risks and use this opportunity to reflect on your choice.

Reflection
Why is this risk a concern for you? Why is this a risk for others you know? What are the potential consequences of this risk?

Your reflection

Reflection on the Risks of Using LLMs The risks of using large language models (LLMs) are concerning to me because of my deep interest in **data science, AI ethics, and quantum computing**. As someone who works with **large datasets, AI models, and prompt engineering**, I understand how biases, misinformation, and security vulnerabilities can impact the reliability and fairness of AI-driven decisions. These risks are not just theoretical but have real-world implications, especially in fields like **healthcare, finance, education, and governance**. **Why This Is a Risk for Me** 1. **Data Privacy Concerns** – Since LLMs often rely on vast amounts of data, including personal or proprietary information, improper handling can lead to **data leaks or misuse**. This is crucial when working with **confidential enterprise data or research datasets**. 2. **Bias in AI Outputs** – LLMs can reinforce **societal biases**, leading to unfair or discriminatory outcomes. As an aspiring researcher, I want AI to be **fair and explainable**, which makes bias a critical issue. 3. **Misinformation & Hallucinations** – LLMs can generate **factually incorrect** or misleading content. In **academic research, Ph.D. work, and AI development**, misinformation can compromise the credibility of findings and decisions. #### **Why This Is a Risk for Others** 1. **Security Threats** – LLMs can be exploited for **cyberattacks, phishing, and fraud**. Companies relying on AI for customer support, automation, or decision-making might become vulnerable to **data breaches**. 2. **Overreliance & Skill Degradation** – Dependence on AI tools may **reduce critical thinking skills** in students, professionals, and even researchers, leading to **lower-quality human oversight**. 3. **Manipulation & Deepfakes** – AI can be misused to create **convincing fake content**, affecting **elections, reputations, and public trust**. #### **Potential Consequences** - **For individuals** – Privacy loss, misinformation spread, reduced analytical skills. - **For

businesses** – Legal liability, financial losses due to biased or incorrect AI-driven decisions. - **For society** – Erosion of trust in AI, misinformation crises, increased inequality due to biased AI models. ### **Final Thought** While LLMs offer **powerful benefits**, these risks emphasize the need for **robust AI governance, ethical AI development, and strict security measures**. AI should be a tool to **augment human intelligence, not replace it recklessly**.

Things to think about
Great! It's good to know that the model is aware of this risk.

Choose a risk not suggested by the LLM
Is there a risk not listed by the LLM that you can think of? If so, please take this moment to reflect on it.

Reflection
Why is this risk a concern for you? Why do you think the LLM did not list this risk in its output?

Enter your response here, there's no right or wrong answer

Submit

Exercise End

Great work! Like any other powerful technology, we need to make sure we use it responsibly and in ways that align with our values. Understanding the risks is an important part to being able to make informed decisions about when and how to use Generative AI.