AI Career Exploration and Skill Gap Analysis

Ali Muhammad Asad
aa07190

1. Job Description

Job Title: AI Safety Engineer - Link can be found here.
Location: Bay Area - San Francisco and Palo Alto
Expires: 04/30/2026
xAI is dedicated to creating AI systems that can accurately understand the universe and aid humanity in its pursuit of knowledge. The AI safety engineer role focuses on identifying safety shortcomings in their AI system, Grok, and implementing methods from machine learning safety literature to enhance its safety. This involves designing evaluations to assess potential risks, collaborating with other teams to integrate safety measures, and innovating new ideas to develop AI systems that can accurately understand the universe.

2. Skills Breakdown

Tech Stack:

- Python

- JAX and XLA

- Rust

- Cuda (C++ and Triton)

Key Responsibilities:

- Identify shortcomings in Grok

- Adapt existing or design new evaluations to assess potential risks

- Implement methods from ML safety literature to improve Grok's safety

- Collaborate with other teams to ensure safety measures are integrated into the development pipeline

- Innovate new ideas to advance the goal of developing AI systems that understand the universe

Required Skills and Qualifications:

- Experience working on ML safety problems with demonstrated exceptional work

- Expertise in machine learning and fine-tuning large language models

- Experience in developing and managing large-scale machine learning systems

- Ability to keep up with multiple ML safety domains

Soft Skills: Strong communications skills, hands-on approach to contribute directly to the company's mission.

3. Self Assessment

Throughout my degree, I've taken up various courses in AI and machine learning, including Computational Intelligence, Deep Learning, Probabilistic Graphical Models, Reinforcement Learning, Large Language Models, and currently Generative AI: Security, Ethics and Governance. Thus, I have engaged in AI and ML projects, including my Final Year Project which basically is developing a Large Language Model from scratch for a low resource language which would include evaluation, assessment, and safety measures. My projects have also been aligned with the company's mission of developing AI systems that can accurately understand the universe and aid humanity, such as in the preservation of literature (deep learning), predicting guide efficacy of Crispr-Cas systems (probabilistic graphical models) that can aid biologists in their research for effective gene editing and drug discovery, and developing a large language model for a low resource language, which can aid the people of that region in their communication and understanding of the world. Therefore, I am well versed in Python, however, I do not have experience with JAX and XLA, Rust, or Cuda. I am willing to learn these languages and tools to adapt to the company's tech stack. I have experience in developing and managing large-scale machine learning systems, and I am confident in my ability to adapt to multiple ML safety domains. I have strong communication skills and a hands-on approach to contribute directly to the company's mission.

4. Knowledge & Skills Gap

I do not have prior experience with JAX and XLA, Rust, or Cuda. I will need to learn these languages and tools to adapt to the company's tech stack. I will also need to gain more experience in ML safety problems and fine-tuning large language models to meet the company's requirements - which I will have hopefully done by the end of my Generative AI course, and towards the conclusion of my Final Year Project. I am also not familiar with Triton, and will need to learn this tool to adapt to the company's tech stack. Grok is also something I am not familiar with, and I will need to learn about this system to identify its shortcomings and implement methods from ML safety literature to improve its safety.

5. Action Plan for Filling the Gaps

To learn the above-mentioned languages and tools, I will take online courses and tutorials, and work on projects that require the use of these languages and tools. I will also read research papers and articles on ML safety problems and fine-tuning large language models to gain more experience in these areas. I will also read up on Triton and Grok to familiarize myself with these systems.

6. Reflection

The exercise to find an AI job description and analyze my skills and knowledge gap has been very insightful. It has helped me identify areas where I need to improve and develop my skills further. I am excited to take on the challenge of learning new languages and tools, and gaining more experience in ML safety problems, something which I have not done, neither considered

before. I am confident that I will be able to fill the gaps in my knowledge and skills and be able to apply for the job in the future.