56

ML Safety Scholars Summer 2022 Retrospective

by **TW123** Nov 1 2022

AI safety | Career choice | Community | AI alignment | Fellowships and internships

Show all topics

TLDR

This is a report on the Machine Learning Safety Scholars program°, organized over the summer by the Center for AI Safety. 63 students graduated from MLSS: the list of graduates and final projects is here. The program was intense, time-consuming, and at times very difficult, so graduation from the program is a significant accomplishment.

Overall, I think the program went quite well and that many students have noticeably accelerated their AI safety careers. There are certainly many areas of improvement that could be made for a future iteration, and many are detailed here. We plan to conduct followup surveys to determine the longer-run effects of the program.

This post contains three main sections:

- This TLDR, which is meant for people who just want to know what this document is and see our graduates list.
- The executive summary, which includes a high-level overview of MLSS. This might be of interest to students considering doing MLSS in the future or anyone else interested in MLSS.
- The full report, which was mainly written for future MLSS organizers, but I'm publishing here because it might be useful to others running similar programs.

The report was written by Thomas Woodside, the project manager for MLSS. "I" refers to Thomas, and does not necessarily represent the opinion of the Center for AI Safety, its CEO Dan Hendrycks, or any of our funders.

Visual Executive Summary

MLSS Overview

MLSS was a summer program for mostly undergraduate students aimed to teach the foundations of machine learning, deep learning, and ML safety. The program ended up being ten weeks long and included an optional final project. It incorporated office hours, discussion sections, speaker events, conceptual readings, paper readings, written assignments, and programming assignments. You can see our full curriculum here.