IEEE Southeastern Michigan
Presents Distinguished Speaker Talk on:
Computational Safety for Generative AI

IEEE Southeastern Michigan
Electrical and Electronic Engineers Creating Our Future

Large language models (LLMs) and Generative AI (GenAI) are at the forefront of frontier AI research and technology. With their rapidly increasing popularity and availability, challenges and concerns about their misuse and safety risks are becoming more prominent than ever. In this talk, we introduce a unified computational framework for evaluating and improving a wide range of safety challenges in generative AI. Specifically, we will show new tools and insights to explore and mitigate the safety and robustness risks associated with state-of-the-art LLMs and GenAI models, including (i) safety risks in fine-tuning LLMs, (ii) LLM red-teaming and jailbreak mitigation, (iii) prompt engineering for safety debugging, and (iv) robust detection of AI-generated content.

Quick Summary

- When:
  Date: May 18th, 2026
  Time: 06:00 – 7:00 PM (EST/EDT)
- Where:
  Online via Webex (to be shared only after you have a confirmed registration)
- Audience: OPEN to ALL*

Sponsored by IEEE Southeastern Michigan Signals/Computer & Education Society Technical Chapters

***Pre-Registration Required!**

https://events.vtools.ieee.org/m/559586

IEEE Signal Processing Society — CELEBRATING 75 YEARS

IEEE COMPUTER SOCIETY

IEEE Education Society

IEEE Southeastern Michigan Section