I. DAVID REIN

817-800-0794 | idavidrein@gmail.com | idavidrein.github.io | linkedin.com/in/idavidrein | github.com/idavidrein

WORK & RESEARCH EXPERIENCE

Model Evaluations and Threat Research (METR): ML Research Scientist   August 2024 - Present
· **Leading the General Autonomy team**, developing the science of measuring the capabilities of AI systems.

NYU: Research Scientist (Supervised by Sam Bowman)   September 2022 - Present
· Research focused on scalable oversight methods and evaluations
· *GPQA: A Graduate-Level Google-Proof Q&A Benchmark:* A new benchmark that has quickly become a standard evaluation for new foundation models (e.g. Claude 3, Llama 3, Gemini 1.5 Pro, Reka Flash)
· *Debate Helps Supervise Unreliable Experts*: First positive results for AI Safety via Debate in a realistic setting.

Cohere: Member of Technical Staff   September 2020 - November 2023
· Early employee, **co-founded the Embeddings team**, scaled up contrastive learning methods
· Trained large production LLMs on TPU v4 pods with Jax, designed and implemented core infrastructure.

Duke University   January - August 2020
· *Classification with Strategically Withheld Data:* AAAI 2021
· Evaluated the performance of a novel classification algorithm that is provably robust to strategically withheld data.
· Designed the graduate-level *Foundations of Deep Learning* (STA 790) course with Dr. David Banks

EDUCATION

Duke University   December 2021
· B.S. Computer Science, B.A. Philosophy **-** GPA: 3.81/4.0
· *Machine Learning (graduate level), Bayesian and Modern Statistics (graduate level), Applied Ethics, Metaphysics, Ethics and AI, Computational Microeconomics, Algorithms, Operating Systems, Computer Architecture, Science and Social Justice*

LEADERSHIP

Duke Undergraduate Machine Learning: Co-President   August 2018 - 2020
· Lead organizer of the 2019 Duke Datathon with 350+ participants; raised $20k+ in sponsorship for the event.
· Hosted ~20 speakers from leading industry and research labs for seminars and workshops; average 20-40 attendees.

Duke Effective Altruism (EA): Co-President   August 2019 - 2021
· Led and helped design the Arete Fellowship, a 12-week discussion-based program to introduce 20+ undergrads to EA.

ACTIVITIES AND SKILLS

Reinforcement Learning Implementations   May - June 2019
· Implemented REINFORCE (VPG), A2C with Generalized Advantage Estimation, and Proximal Policy Optimization.

ASA Duke DataFest: 1st Place - Best Insight Award   April 2019
· Competed in a group against 425+ undergraduate (2/3) and graduate (1/3) students from 8 universities.
· Predicted fatigue from biometric data of the Canada women's rugby 7s team with a Cox proportional hazards model.

Kenan Institute for Ethics Policy Prize in the Ethics of Emerging Tech: 2nd Place   April 2019
· Co-authored research paper on mechanics, ethics, and international policy of orbital debris and anti-satellite weaponry
· Presented paper at the 2019 Duke Conference on the Ethics of Emerging Technology.

Languages and Tools: Python, NumPy, TensorFlow, Jax, PyTorch, Scikit-Learn, Pandas