Jonathan Liu

jonathanliu@princeton.edu | +1(781)-694-0117 | https://github.com/liujonathan24/

EDUCATION

Princeton University **May 2028**

A.B. Computer Science (GPA: 3.9, Major GPA: 4.0)

Relevant Coursework: Systems & ML; Natural Language Processing; (Graduate) Reinforcement Learning; Operating Systems; Data Structures and Algorithms; Probability, Stochastic Systems; Multivariate Analysis; Linear Algebra

MACHINE LEARNING RESEARCH

Incoming Research Scientist (PhD Role) — Abridge AI **Summer 2026**

Research Assistant — Prof. Zhuang Liu Group **Aug 2025 - Present**

- Fine-tuned LLMs for paper-acceptance prediction using LlamaFactory and TRL with Deepspeed for tensor parallelism and VLLM for inference. Torch FSDP2 is used for distributed training. Publication submitted to COLM.
- Analyzing multimodal LLM behavior relative to ablated architectures to identify key design contributions using Jax.

Machine Learning Research Intern — BBN Technologies **Jun 2025 - Aug 2025**

- Demonstrated that using an LLM to respond to and evaluate answers of a question dataset may produce statistically significant, but non-generalizable results. First author on paper presented at 2025 NeurIPS workshop GenAI4Health.
- Detected hallucinations in clinical LLMs using agentic RAG and LLM-as-a-judge pipelines (Autogen, LangChain, FAISS, Qwen embeddings).

Research Assistant — Princeton Computer Vision Lab (Prof. Jia Deng) **Aug 2024 - Present**

- Coauthor on paper *Infinigen-Sim: Procedural Generation of Articulated Simulation Assets* (accepted to CoRL LSRL), which contributes a framework for the procedural generation of *articulated* simulation assets, improving success rate of 0-shot sim2real tasks by 70%. Optimized Genesis and MuJoCo to reduce instability in complex scenes using convex hull deconstruction techniques implemented in C.

PROJECTS AND EXTRACURRICULARS:

Sight Support — Personal Project **Spring 2020 - Winter 2025**

- Assistive technology app for blind individuals that detects social cues in online meeting using multi-modal LLMs. App uses Faster-Whisper and LM Studio. Includes prior research on the use of prompting techniques, such as one-/multi-shot and chain-/tree-/graph-of-thought, to seek explainability in Facial Expression Recognition (FER) and the development of “earcons,” short audio bytes, that represent social cues.
- Awards: Best Poster Award Winner (4 selected / 60) at MASC-SLL 2025, Presenter at 2024 Yale Engineering and Science symposium (60 selected / 50,000 students), 1st place winner ($2500) at 2024 Amazon, Perkin School for the Blind, and MIT’s #HackDisability.

Discovering Transformer Circuits with Edge Pruning — NLP Course Project **Spring 2025**

- Replicated the "Edge Pruning" mechanistic interpretability method on GPT-2, identifying sparse computational subgraphs for standard tasks, and extended methods to identify computational subgraphs for the induction mechanism.

RL for the Andrews-Curtis Conjecture — Personal Project **Spring 2025 - Present**

- Researching the application of reinforcement learning to long-horizon tasks, namely solving the the Andrews-Curtis Conjecture, using A3C, PPO, and contrastive learning algorithms.
- Award: Accepted to MAA MathFest 2025 Student Poster Session.

AWARDS

- Qualifier to 2025 International Collegiate Programming Contest (ICPC) Regional Round
- 4x AIME qualifier, top 40% Putnam Competition

SKILLS AND INTERESTS

- Attend and present at Princeton’s Computer Vision and Natural Language Processing reading groups.
- **Technical Skills:** Python (PyTorch, Jax, Pandas, NumPy, Hydra, SciPy, Gym); C; Java; Linux; Slurm; Bash; Triton.
- **Interests:** Classical cello (Princeton University Orchestra); running (Jersey City Marathon ’25); video games