

## Education

- Massachusetts Institute of Technology** Cambridge, MA  
*BS & MEng, Computer Science & Cognitive Science* Sep 2021 – May 2025
  - Coursework:** Algorithms I & II, Machine Learning, Deep Learning, TinyML, Linear Algebra, Probability, Computational Cogsci, AI Ethics, Game Theory, Hardware for Deep Learning
  - Master's Thesis:** Towards a Spectral Understanding of Language Model Fine-Tuning

## Experience

- Y Combinator** San Francisco, CA  
*Summer Fellow* June 2025 – Aug 2025
  - Received \$20,000 equity-free grant. Working on RL for LLMs.
- MIT CSAIL** Cambridge, MA  
*Language & Intelligence Group Researcher* Feb 2024 – May 2025
  - Studied LLM fine-tuning. Published paper currently under review.
- Google DeepMind** New York City, NY  
*Research Engineering Intern* Jun 2024 – Aug 2024
  - Implemented and ran hundreds of experiments across thousands of TPUs to measure and improve Gemini's factuality, especially in multi-modal contexts.
  - Aligned state-of-the-art factuality auto-classifier with human factuality labels.
- Cleanlab** San Francisco, CA  
*Machine Learning Engineering Intern* Jan 2024 – Mar 2024
  - Developed and implemented novel ways to detect data issues in order improve data quality.
  - Wrote in production code to detect low quality text with high precision.
- Numenta** Redwood City, CA  
*Software/Machine Learning Engineering Intern* May 2023 – Aug 2023
  - Created novel PEFT fine-tuning methods for LLMs to meet strict customer and hardware constraints.
  - Wrote code to support efficient sparse neural networks.
- MIT CSAIL** Cambridge, MA  
*Undergraduate Researcher* Dec 2021 – May 2023
  - Studied LLMs and their use cases. Published separate papers in NeurIPS FMDM '22(100+ citations) and PNAS '22(150+ citations).

## Selected Work

- Exploring the causes & effects of quantization-induced degradation in LLMs** (see [blog](#))
  - Examined why QiD occurs, and which transformer layer type causes it.
- Exploring Activation-aware Quantization for LLMs** ([reeceshuttle.me/assets/6\\_5940\\_Final\\_Project\\_Report.pdf](https://reeceshuttle.me/assets/6_5940_Final_Project_Report.pdf))
  - Studied mixed precision and AWQ to extend lossless quantization past 8-bit.
- Analyzing Inference Optimizations for Transformers** ([reeceshuttle.me/assets/6.5930\\_Project.pdf](https://reeceshuttle.me/assets/6.5930_Project.pdf))
  - Studied inference optimizations in the attention module of transformers.
- Sparsity in Transformers** ([github.com/reeceshuttle/958](https://github.com/reeceshuttle/958))
  - Systematically measured the sparsity of weights and attention scores across several transformer models.
- MIT Pokerbots** ([github.com/reeceshuttle/poker-bot](https://github.com/reeceshuttle/poker-bot))
  - Placed in the top 10% of entries in 2023 MIT Pokerbots competition and awarded a cash prize.
- Gabor filter-constrained CNNs** ([github.com/samacqua/gabor-constrained-nns](https://github.com/samacqua/gabor-constrained-nns))
  - Trained unique Convolutional Neural Networks by seeking inspiration from the human brain.
- PyTorch, but in NumPy** ([github.com/reeceshuttle/numpytorch](https://github.com/reeceshuttle/numpytorch))
  - Implemented basic PyTorch functionality using only NumPy arrays.

## Technical Skills & Interests

- Languages:** Python, C, C++, HTML/CSS, JavaScript, Julia, LaTeX, RISC-V, R
- Tools/Frameworks:** PyTorch, JAX, Git, Docker, WandB, AWS
- Interests:** AI, neuroscience, reading, hiking, aviation, space flight, history