brando9@stanford.edu
Stanford, CA, 94305.

Brando Miranda

Gates Bldg., 353 Jane Stanford Way

Google Scholar; Website; Stanford Profile; MIT Website

Research Interests

Data-centric machine learning for Frontier Models (FMs), Artificial General Intelligence (AGI), and reasoning for mathematics and verified code. Focus on scaling laws, formal reasoning, and the emergence of capabilities in large language models.

Education

Ph.D. in Computer Science *2022-2026 (expected)*
Stanford University GPA: 4.045/4.0
Advisor: Prof. Sanmi Koyejo, Stanford Trustworthy AI Research (STAIR) group
Fellowships: Stanford School of Engineering Fellowship, EDGE Scholar

Master of Engineering in Electrical Engineering and Computer Science *2014-2016*
Massachusetts Institute of Technology GPA: 4.8/5.0
Advisor: Prof. Tomaso Poggio, Center for Brains, Minds and Machines (CBMM)
Focus: Deep learning theory

Bachelor of Science, Computer Science and Engineering *2010-2014*
Massachusetts Institute of Technology minor in Mathematics & Music

Awards & Honors

- **NeurIPS Outstanding Main Track Paper Award** (December 2023) (top 0.4%, only 2 papers selected)
- **ICML Outstanding Paper TiFA Workshop Award** (July 2024)
- **EDGE Scholar**, Stanford University (September 2022)
- **Stanford School of Engineering Fellowship** (September 2022)
- **Honorable Mention**, Ford Foundation Fellowship (2020, 2021)
- **Best Research Project Award**, UIUC graduate course CS 598 "Learning to Learn" (December 2020)
- **Most Cited Paper Certificate**, International Journal of Automation & Computing (IJAC) (December 2019)
- **Computer Science Excellence Saburo Muroga Endowed Fellow**, UIUC (2019-2020)
- **Sloan Scholar**, Alfred P. Sloan Foundation's Minority Ph.D. (MPHD) Program (2018-2019)
- **Grainger Engineering SURGE Fellowship** (2018-2019)

Selected Publications

VeriBench: End-to-End Formal Verification Benchmark for AI Code Generation in Lean 4 (2025)
Brando Miranda, Zhangir Zhou, Anima Nie, Elio Obbad, Leni Aniva, Kai Fronsdal, William Kirk, Dogus Soylu, et al.
2nd AI for Math Workshop @ ICML 2025

CoDaPO: Confidence and Difficulty-Adaptive Policy Optimization for Post-Training Language Models (2025)
Zhangir Zhou, Xingrun Lu, Cheng Cao, Brando Miranda, Tong Liu, Bo Han, Sanmi Koyejo
2nd AI for Math Workshop @ ICML 2025

Why Has Predicting Downstream Capabilities of Frontier AI Models with Scale Remained Elusive? (2025)
Rylan Schaeffer, Hailey Schoelkopf, Brando Miranda, Gabriel Mukobi, Varun Madan, Adam Ibrahim, Herbie Bradley, Stella Biderman, Sanmi Koyejo
International Conference on Machine Learning (ICML) 2025 & **ICML TiFA Workshop Outstanding Paper Award 2024**
[arXiv]

Putnam-AXIOM: A Functional & Static Benchmark for Measuring Higher Level Mathematical Reasoning in LLMs (2025)
Aryan Gulati, Brando Miranda, Eric Chen, Emily Xia, Kai Fronsdal, Bruno de Moraes Dumont, Sanmi Koyejo
International Conference on Machine Learning (ICML) 2025 & **NeurIPS MATH-AI Workshop 2024**

Pantograph: A machine-to-machine interaction interface for advanced theorem proving, high level reasoning, and data extraction in Lean 4 (2025)
Leni Aniva, Chuyue Sun, Brando Miranda, Clark Barrett, Sanmi Koyejo
International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS) 2025
[Springer Link]

Causally Quantifying the Effect of Test Set Contamination on Generative Benchmarks (2025)
Rylan Schaeffer, Brando Miranda, Jiaqi Kazdan, Kaivu Liu, Amir M. Ahmed, Niloofar Mireshghallah, et al.
NeurIPS 2025 Workshop on Evaluating the Evolving LLM Lifecycle

Position: Machine Learning Conferences Should Establish a 'Refutations and Critiques' Track (2025)
Rylan Schaeffer, Jiaqi Kazdan, Yarin Denisov-Blanch, Brando Miranda, Max Gerstgrasser, et al.
Preprint 2025

Failures to Find Transferable Image Jailbreaks Between Vision-Language Models (2024)
Rylan Schaeffer, Dan Valentine, Luke Bailey, James Chua, et al., Brando Miranda, et al., Sanmi Koyejo, Ethan Perez
International Conference on Learning Representations (ICLR) 2024 & **NeurIPS Red Teaming GenAI Workshop 2024**
[OpenReview]

Does Maximizing Neural Regression Scores Teach Us About The Brain? (2024)
Rylan Schaeffer, Mikail Khona, Sankarshan Chandra, Michael Ostrow, Brando Miranda, Sanmi Koyejo
UniReps Workshop 2nd Edition, NeurIPS 2024

Quantifying the Importance of Data Alignment in Downstream Model Performance (2024)
Keshav Chawla, Aditya Sahai, Marco DePavia, Sathvik Sundar, Brando Miranda
ICLR 2024 Workshop on Data-centric Machine Learning Research (DMLR)

An Evaluation Benchmark for Autoformalization in Lean4 (2024)
Aryan Gulati, Devanshu Ladsaria, Shubham Mishra, Jasdeep Sidhu, Brando Miranda
The Second Tiny Papers Track at ICLR 2024

Are Emergent Abilities of Large Language Models a Mirage? (2023)
Rylan Schaeffer, Brando Miranda, Sanmi Koyejo
Neural Information Processing Systems (NeurIPS) Outstanding Main Track Paper Award & Oral 2023
Featured in: New York Times, Quanta Magazine, Forbes, Stanford HAI
Cited in the 2024 Economic Report of the President (White House)
[arXiv] [OpenReview]

Morph Prover v0 7b: The 1st Frontier Model for the Lean 4 Formal Verification Programming Language (2023)
Brando Miranda (Machine Learning Research Scientist Consultant) & Morph Labs Team
[Blog] [Hugging Face Model Card]

Beyond Scale: the Diversity Coefficient as a Data Quality Metric Demonstrates LLMs are Pre-trained on Formally Diverse Data (2023)
Alycia Lee, Brando Miranda*, Patrick Yu, Oluwasanmi Koyejo* (*equal contribution)
ICML Data-Centric Machine Learning Workshop & ICML Deployable Generative AI Workshop 2023
[arXiv] [Code]

Why and when can deep-but not shallow-networks avoid the curse of dimensionality: a review (2017)
Tomaso Poggio, Hrushikesh Mhaskar, Lorenzo Rosasco, Brando Miranda, Qianli Liao
International Journal of Automation and Computing 2017, Most Cited Paper Certificate
[Journal Link]

Google Scholar metrics: 2,948+ citations — h-index: 17 — i10-index: 22

Research & Professional Experience

Stanford University - Stanford, CA *September 2022 - 2026 (expected)*
Ph.D. Student in Computer Science

- Conducting research on data-centric machine learning for Frontier Models, formal reasoning and verification, and scaling laws for LLMs
- Published award-winning work challenging prevailing notions of "emergent abilities" in large language models
- Leading development of VeriBench and Putnam-AXIOM benchmarks for evaluating AI mathematical reasoning
- Research mentor for undergraduate and master's students focused on data quality for Foundation Models

Morph Labs - Remote *2023*
Machine Learning Research Scientist Consultant

- Made key contributions to Morph Prover v0-7B, the first frontier model for Lean 4 formal verification language
- Led the creation of the embedding system for Moogle.ai, the first search engine for verified code in Lean

Wise Agents - Stanford Spin-out *2023*
AI Research Consultant

- Consulted on AI agent-based systems to transform sales performance
- Company featured in Forbes Mexico

IBM Research - Yorktown Heights, NY *May 2022 - August 2022*
Graduate Research Intern

- Developed transformer models to predict mathematical formal proofs using the Coq theorem proving language

University of Illinois Urbana-Champaign - Urbana-Champaign, IL *September 2018 - 2021*
Ph.D. Student in Computer Science

- Conducted foundational research on meta-learning, transfer learning, and data diversity
- Published multiple papers on the relationship between task diversity and learning performance

IBM Research - Yorktown Heights, NY *May 2021 - August 2021*
Graduate Research Intern

- Studied optimization and architectural inductive biases of Transformer models for program type inference

MIT CBMM (Center for Brain Minds & Machines) - Cambridge, MA *June 2015 - September 2018*
Research Assistant

- Conducted deep learning theory research with Prof. Tomaso Poggio (who also mentored DeepMind co-founder Demis Hassabis)
- Contributed to highly-cited work on generalization theory and function approximation for neural networks

Media Coverage

- **White House Economic Report of the President (2024)**: Our work on emergent abilities was cited in the official economic report to the President
- **The New York Times (2023)**: "Silicon Valley Confronts the Idea That the 'Singularity' Is Here"
- **Quanta Magazine (2024)**: "How Quickly Do Large Language Models Learn Unexpected Skills?"
- **Forbes (2023)**: "AI 'Emergent Abilities' Are A Mirage, Says AI Researcher"
- **Andrew Ng (2024)**: Endorsed our paper as evidence that AGI development will be smooth and predictable
- **Additional coverage**: Stanford HAI, Y Combinator News, Vice, Medium, HackerNews, NeurIPS blog

Teaching Experience

University of Illinois Urbana-Champaign *August 2020 - December 2020*
Graduate Teaching Assistant

- CS 446 Machine Learning
- Designed problem sets and exams; held weekly office hours

Massachusetts Institute of Technology *2014-2016*
Graduate Teaching Assistant

- Statistical Learning Theory & Applications (9.520/6.860) with Prof. Lorenzo Rosasco & Tomaso Poggio
- Introduction to Algorithms (6.006) with Prof. Nancy Lynch, Bruce Tidor and Aleksander Madry
- Design & Analysis of Algorithms (6.046) with Prof. Shafi Goldwasser, Dana Moshkovitz, Nir Shavit
- Introduction to Machine Learning (6.036) with Prof. Tommi Jaakkola, Suvrit Sra & Regina Barzilay

Leadership & Service

Research Mentorship *2016 - Present*

- Mentored undergraduate and graduate students at Stanford, UIUC, and MIT CBMM
- Led research teams on data quality metrics, meta-learning, and formal reasoning

Academic Service *2018 - Present*

- Reviewer for ICLR 2020, JMLR 2018
- Graduate advisor for Latinos in Computer Science (LCS) at UIUC
- Founding member of Stanford AI for Lean, a research community advancing AI for Lean theorem proving and formalizing mathematics
- Founded "Stanford Bachata Sensual & Brazilian Zouk" and "UIUC Bachata Sensual & Zouk" official student organizations

Outreach *2017 - Present*

- Distributed Research Experiences for Undergraduates (DREU) UIUC 2019
- Undergraduate Research Opportunity Program (UROP) MIT 2017-2018
- Engineering of Intelligence Team (EIT) CBMM MIT 2017-2018

Skills & Tools

Programming: Python, PyTorch, TensorFlow, JAX, Lean 4, OCaml, Coq

ML/AI Tools: Hugging Face, Weights & Biases, DSPy, git, UNIX

Languages: English, Spanish (native fluency)