

Lindsey Ma

Email: lm3879@columbia.edu

GitHub: github.com/Lindsey-cyber

Portfolio: lindseyma.com

Education

Columbia University

Sept 2024 – May 2026

M.S. in Statistics

Relevant Coursework: Deep Learning, Advanced Machine Learning, Statistical Analysis of Neural Data

Renmin University of China

Sept 2019 – Jun 2023

B.A. in Economics

Technical Skills

Languages: Python, SQL, R

ML: PyTorch, Hugging Face, scikit-learn

Tools: Linux, Docker, Neo4j, vLLM, FastAPI

Research Experience

Research Assistant - Multimodal Learning & Knowledge-Guided Reasoning *May 2026 – Present*
Columbia University, Department of Statistics

- Probed **intermediate representations of pretrained vision models** to study how plant morphology emerges in learned feature spaces; analyzed DINOv3 patch embeddings with PCA and K-Means and developed downstream classifiers/probes to quantify morphological information across representations.
- Built an end-to-end **VLM + knowledge graph + LLM reasoning** system for fine-grained species classification, improving top-1 accuracy from **60% to 86%** on an iNaturalist-derived benchmark spanning 100+ species.
- Designed controlled ablations with correct, shuffled, and corrupted biological relations to isolate the contribution of external knowledge; ontology refinement produced an additional **6% improvement** in top-1 accuracy.

Research Assistant - Self-Supervised Learning for Neural Imaging

Jan 2025 – May 2026

Center for Theoretical Neuroscience, Columbia Zuckerman Institute

- Designed and trained a **self-supervised blind-spot neural network in PyTorch** for denoising two-photon calcium-imaging movies under temporally correlated noise.
- Conducted systematic ablations over masking strategies, receptive fields, and preprocessing across five real and simulated imaging datasets, evaluating latent-signal recovery and neuronal morphology preservation.
- Rebuilt the lab's PMD preprocessing pipeline, reducing residual video energy by **20%** and downstream decoding variance by **7%** versus baseline.

Selected Project

FlyWire Connectome Analysis - Computational Statistics Project

May 2025

- Applied **Bayesian stochastic block models** to the FlyWire whole-brain connectome to infer cell-type organization; evaluated community structure with posterior predictive checks and visualized graph representations with UMAP.

Additional Experience

Columbia AI Alignment Club - Technical Fellow

Feb 2025 - Present

- Technical fellowship focused on **frontier-model evaluation and mechanistic interpretability**, including activation steering and latent-knowledge elicitation.

Hackathon Team Lead - 2nd Place, XRPL (\$5K); 3rd Place, Stellar (\$3K) *Sept 2025 – Oct 2025*

- Led cross-functional teams to ship two production-style applications in 24-hour blockchain hackathons; owned backend architecture, transaction pipelines, integrations, and end-to-end testing.