Yuan Pu

Email: yuanpu.00@gmail.com | Homepage: yuan-pu.github.io | Updated: July 2026

Education

Duke University August 2025 – May 2030 (expected)
Doctor of Philosophy Program in Computer Science *Durham, NC*

- GPA: 4.0/4.0

Brown University August 2019 – May 2023
Bachelor of Science in Computational Biology (Computer Science track) *Providence, RI*

- GPA: 4.0/4.0; magna cum laude; Computational Biology Departmental Honors

Skills

Programming Languages: Python, R, MATLAB

Frameworks: PyTorch, TensorFlow, HuggingFace Transformers, NumPy, pandas, stan

Tools & Platforms: Git, Jupyter, OpenAI/Anthropic APIs, Weights & Biases, Google Cloud Platform, HPC clusters

Methods: LLM Evaluation, AI-Human Interactions and Collaborations, Time Series Modeling

Selected Publications (*co-first authorship)

[C1] **MedRedFlag: Investigating how LLMs Redirect Misconceptions in Real-World Health Communication** | *ACL 2026 (findings), SAIL 2026*

Sraavya Sambara*, Yuan Pu*, Ayman Ali, Vishala Mishra, Lionel Wong, Monica Agrawal

[C2] **Trajectory Flow Matching with Applications to Clinical Time Series Modelling** | *NeurIPS 2024 (spotlight)*

Xi Zhang*, Yuan Pu*, Yuki Kawamura, Andrew Loza, Yoshua Bengio, Dennis Shung, Alexander Tong

[C3] **Human-Algorithmic Interaction Using a Large Language Model-Augmented Artificial Intelligence Clinical Decision Support System** | *CHI 2024*

Niroop Channa Rajashekar*, Yeo Eun Shin*, Yuan Pu*, Sunny Chung, Kisung You, Mauro Giuffrè, Colleen Chan, Theo Saarinen, Allen Hsiao, Jasjeet Sekhon, Ambrose Wong, Leigh Evans, René Kizilcec, Loren Laine, Terika McCall, Dennis Shung

[C4] **Counting Clues: A Lightweight Probabilistic Baseline Can Match an LLM** | *ML4H 2025 (findings)*

Furong Jia*, Yuan Pu*, Finn Guo, Monica Agrawal

[J1] **Assessing the Usability of GutGPT: A Simulation Study of an AI Clinical Decision Support System for Gastrointestinal Bleeding Risk** | *npj Digital Medicine 2025*

Sunny Chung, Mauro Giuffrè, Niroop Rajashekar, Yuan Pu, Yeo Eun Shin, Simone Kresevic, Colleen Chan, Shinpei Nakamura-Sakai, Kisung You, Theo Saarinen, Allen Hsiao, Ambrose H Wong, Leigh Evans, Terika McCall, Rene Kizilcec, Jasjeet Sekhon, Loren Laine, Dennis Shung

Research Experience

Graduate Student Aug 2025 – Present
Duke University *Durham, NC*

Evaluating LLM in Real-World Patient Health Communication | *advised by Monica Agrawal*

- Co-developed `MedRedFlag`, a benchmark of 1,100+ real-world patient health questions containing problematic assumptions, curated via a semi-automated LLM-based annotation pipeline. Designed LLM-as-judge evaluation metrics to assess how LLMs reply to questions with medical misconceptions; found systematic gaps between state-of-the-art LLMs and physician behavior. (ACL 2026 findings and SAIL 2026, co-first author)
- Conducting follow-up studies investigating the safety, robustness, and real-world utility of LLM-generated health advice, as well as how consumers perceive and respond to such advice.

Probing LLM Reasoning with Probabilistic Baselines | *advised by Monica Agrawal*

- Co-designed a lightweight Naive Bayes method over concept-diagnosis co-occurrence statistics from LLM pretraining corpora. Demonstrated that FBPR matches 7B LLM accuracy on MedQA diagnosis benchmarks with complementary error patterns, clarifying the role of corpus frequency in LLM clinical reasoning. (ML4H 2025, co-first author)

Postgraduate Associate Aug 2023 – May 2025
Yale School of Medicine *New Haven, CT*

Generative Modeling for Time Series with Flow Matching | *advised by Dennis Shung and Alex Tong*

- Designed and implemented Trajectory Flow Matching in PyTorch, a scalable generative framework for training Neural SDEs on irregularly-sampled clinical time series, with modules for uncertainty quantification, interval prediction, and conditional generation (NeurIPS 2024 spotlight, co-first author, top 3%)
- Constructed and cleaned clinical time series benchmarks and conducted large scale experiments. Achieved state-of-the-art results, outperforming Neural ODE/SDE and Latent ODE baselines.

Evaluating Usability and Trust in an LLM-Augmented Decision Support System | *advised by Dennis Shung*

- Co-led evaluation of an LLM-augmented clinical decision support system, designing a mixed-methods study with clinician participants across simulated emergency scenarios to measure usability, trust calibration, and decision-making with AI assistance. (CHI 2024, co-first author)
- Led quantitative analysis of survey data, establishing empirical baselines for clinician engagement patterns with LLM-augmented CDSS in emergency settings. Conducted thematic analysis of clinician interviews, identifying key themes in perceived usability, trust calibration, and barriers to LLM-assisted clinical decision-making.

Large-Scale Clinical Cohort Data Analysis | *advised by Dennis Shung and Darrick Li*

- Engineered a data extraction and analysis pipeline over a 670k patient clinical database using Python. Applied survival analysis and risk factor characterization for post-intervention GI bleeding. (journal submission under review, co-first-author)

Undergraduate Research Assistant — Sept 2020 – May 2023

Brown University — *Providence, RI*

Deep Learning for Gene Expression Prediction from Epigenetic Data | *advised by Ritambhara Singh*

- Built attention-augmented RNN models in Python to predict gene expression from multi-modal epigenetic features across patients. Conducted cross-patient transfer analysis on glioblastoma stem cell data to identify predictive input features.

Survival Analysis and Biomarker Discovery in Cancer Genomics | *advised by Mamiko Yajima*

- Analyzed clinical records and genomic profiles of AML patients, applying survival analysis and differential expression methods to identify prognostic biomarkers and enriched pathways.

Data Science Intern — Jan 2022 – May 2023

Brown University Computational Biology Core — *Providence, RI*

Small RNA-seq Data Pipeline for Toxicogenomics | *advised by Daniel Spade and August Guang*

- Built and executed a small RNA-seq analysis pipeline to identify dosage-dependent expression changes across experimental conditions, performing end-to-end processing from raw sequencing data to pathway-level biological interpretation.

Other Experience

Reviewer for *ML4H*, *CSCW*, *HAI*, *COLM DAIH* — 2025, 2026

Roundtable Junior Chair Clinician-AI Interaction | *ML4H* — 2024

Research Intern AI-powered 3D interior design | *Dymaxion* — Nov 2023 - Aug 2024

Teaching Experience

Teaching Assistant at Brown University

CSCI1430 Computer Vision | *instructed by James Tompkin* — Spring 2022, 2023

CSCI1810/2810 Computational Molecular Biology | *instructed by Sorin Istrail* — Fall 2022

Awards and Honors

Duke University Summer Research Fellowship — 2026

NeurIPS Scholar Award $2400 grant covering conference registration and hotel accommodations — 2024

Brown University Magna Cum Laude — 2023

Brown University Computational Biology Departmental Honors — 2023

Sigma Xi Scientific Honor Society — 2023

Brown SPRINT LINK Program $2000 grant for summer undergraduate research with faculty — 2021