Yuan Pu

Email: yuan.pu@yale.edu | Homepage: yuan-pu.github.io | Updated: November 2024

Education

Brown University Sept 2019 – May 2023

Bachelor of Science in Computational Biology (Computer Science track) - Honors *Providence, RI*

- GPA: 4.0/4.0; magna cum laude; Computational Biology Departmental Honors; Sigma Xi Honors Society
- **Computer Science** coursework in deep learning, machine learning, computational linguistics, computer vision, computational molecular biology, computational population genetics, UIUX
- **Life Sciences** coursework in cell and molecular biology*, genetics, functional genomics, DNA-based techniques*, neurobiology, neural systems, general and organic chemistry* (*with experimental laboratory experience)
- **Math** coursework in calculus, linear algebra, computational probability and statistics, causal inference

Yale University Sept 2023 – Present

Postgraduate Assistant *New Haven, CT*

- Audited courses in fundamentals, recent developments, and applications of Large Language Models (LLMs)

Publications (*equal contribution)

Trajectory Flow Matching with Applications to Clinical Time Series Modelling | Xi Zhang*, Yuan Pu*, Yuki Kawamura, Andrew Loza, Yoshua Bengio, Dennis Shung, Alexander Tong

Neural Information Processing Systems (NeurIPS) 2024 [selected as a spotlight]

Human-Algorithmic Interaction Using a Large Language Model-Augmented Artificial Intelligence Clinical Decision Support System | 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, and Dennis Shung

Computer Human Interaction (CHI) 2024

Assessing the Usability of GutGPT: A Simulation Study of an AI Clinical Decision Support System for Gastrointestinal Bleeding Risk | Colleen Chan, Kisung You, Sunny Chung, Mauro Giuffrè, Theo Saarinen, Niroop Channa Rajashekar, Yuan Pu, Yeo Eun Shin, Loren Laine, Ambrose Wong, René Kizilcec, Jasjeet Sekhon, Dennis Shung

Machine Learning for Health (ML4H) 2023

Machine Learning on Multiple Epigenetic Features Reveals H3K27Ac as a Driver of Gene Expression Prediction Across Patients with Glioblastoma | Yusuke Suita, Hardy Bright, Jr., Yuan Pu, Merih Deniz Toruner, Jordan Idehen, Nikos Tapinos, Ritambhara Singh

Under-review

Research Experience

Postgraduate Associate August 2023 – Present

Yale School of Medicine H+AIM Lab *New Haven, CT*

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

- Developed Trajectory Flow Matching, a novel training method with uncertainty prediction for Stochastic Differential Equations, enhancing scalability and stability in modeling irregularly sampled time series as demonstrated on clinical data.

Evaluation of AI/ML-Enhanced Clinical Decision Support System | *advised by Dennis Shung*

- Collaborated on evaluating usability and trust in human-computer interaction for an LLM-augmented ML clinical decision support system through medical simulation. Collected and conducted analysis on survey results, interview, and conversational statistics between users and the LLM-powered chatbot.

Patient Data Analysis | *advised by Dennis Shung and Darrick Li*

- Performed data extraction and analysis using the Vizient Clinical Data Base, investigating patient characteristics, medication management, and other risk factors associated with percutaneous coronary intervention, post-intervention gastrointestinal bleeding and endoscopy, as well as further clinical outcomes.

Research Intern Nov 2023 – August 2024

Dymaxion *Remote*

AI-powered Interior Design

- Explored LLMs' potential in powering AI generation of indoor furniture arrangement in 2D and 3D with literature review and experiments. Implemented enhancements to an academic research method for commercial application.

Undergraduate Research Assistant May 2022 – May 2023

Brown University Singh Lab *Providence, RI*

Epigenetic Regulation of Gene Expression in GSCs | *advised by Ritambhara Singh*

- Implemented attention-augmented RNNs to predict gene expression from epigenetic data in glioblastoma stem cells (GSCs). Investigated epigenetic regulation of gene transcription in different GSCs by cross-patient analysis.

Data Science Intern Jan 2022 – May 2023

Brown University Computational Biology Core *Providence, RI*

EGME's Impact on Sperm Small RNA Expression | *advised by Daniel Spade and August Guang*

- Processed and conducted analysis on small RNA data in sperm of rats exposed to Ethylene Glycol Monomethyl Ether (EGME). Identified sensitive biomarkers and related biological pathways for EGME's testicular toxicity.

Undergraduate Research Assistant Sept 2020 – Jan 2022

Brown University Yajima Lab *Providence, RI*

Germline Factor DDX4's Role in Cancer | *advised by Mamiko Yajima*

- Revealed the impact of DDX4 expression level on survival in acute myeloid leukemia (AML) patients. Identified co-expressed genes and key biological pathways linked to DDX4, providing insights into its role in AML.

Teaching Experience

Teaching Assistant at Brown University

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

CSCI1810/2810* Computational Molecular Biology (*graduate level) | *instructed by Sorin Istrail* Fall 2022

- Refined contents and solutions of written and coding assignments by monitoring student feedbacks.
- Provided student support for assignments through weekly office hours and online discussions.
- Graded assignments.

Other Experience

Roundtable Junior Chair | *Machine Learning for Health (ML4H) 2024* 2024

Awards and Honors

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

Brown University Magna Cum Laude 2023

Brown University Computational Departmental Honors 2023

Sigma Xi Scientific Honors Society 2023

Hack@Brown 2022 Wolfram Award 2022

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

Skills

Languages: Python, R, MATLAB, Java, JavaScript, HTML/CSS

Frameworks: PyTorch, PyTorch Lightning, TensorFlow, Pandas, Scikit-Learn, React

Tools: Git, Google Cloud Platform, high performance computing clusters