Seyone Chithrananda

seyonec@berkeley.edu | seyonechithrananda.com | Google Scholar | Linkedin

EDUCATION

UC Berkeley Berkeley, CA Bachelor's in Computer Science (Bioengineering minor) *August 2021-May 2025*

- ***Coursework***: Discrete Math & Probability Theory, Convex Optimization, Computational Functional Genomics, Structure and Interpretation of Computer Programs, Designing Information Devices and Systems I & II (linear algebra, circuit analysis, ML), Data Structures, Single + Multivariable Calculus. ***Activities and societies***: Machine Learning at Berkeley, open source developer for DeepChem project (3K stars).

EXPERIENCE

Broad Institute of MIT and Harvard *Research Intern* Cambridge, MA *August 2022-present*

- Visiting student working with Eeshit Dhaval Vaishnav, advised by Prof. Eric Lander (Institute founding director, former WH Science Advisor), studying gene expression programs using single-cell data and variational deep learning methods.

Dyno Therapeutics *ML Research Intern* Cambridge, MA *May 2022-Sept 2022*

- Researching methods for viral protein AAV capsid design, using sequence-to-function graph and sequence-based models (transformers, GNNs, etc). a16z, Google Ventures backed Church Lab startup, with ~$120M Series A (blog post).
- Developed generative structure-to-sequence models to propose high-scoring variant sequences, and examined performance at standard protein redesign, handling epistatic interactions, indels, and at predicting binding on experimentally-validated and in silico fitness landscapes. Preprint accepted at *Machine Learning in Structural Biology* Workshop, NeurIPS 2022.
- Built and grew internal package containing simulated fitness landscapes for benchmarking models on biological sequence design problems. Implemented statistical models for mapping epistatic interactions in progressively rugged landscapes.

Nurix Therapeutics *ML Research Intern* San Francisco, CA *May 2021-September 2021*

- Developing computational strategies for DNA-encoded library design, accounting for multiple sources of experimental variation. Developed graph generative models and genetic algorithms for scaffold-based molecular design, using multi-objective optimization.
- Deployed multiple classification and regression models for screening molecules within core DEL-ML platform. Implemented message-passing, graph convolutional neural networks for binding affinity, ADME-tox modeling.

University of Toronto *Research Intern - Matter Lab (advised by* Alan-Aspuru Guzik) Toronto, ON *Apr. 2020-May 2021*

- Co-developer of SELFIES v1.0, a 100% robust molecular string representation for machine learning models. Developed depth-first graph traversal algorithm and dearomatization code for v1.0 release. Downloaded 9K times to date.
- Published paper at NeurIPS 2020 workshop (lead author) on pre-training strategies for large-scale language modeling of molecules. Implemented transformer models and tokenizers, led integration of NLP-style models into library (600,000 model API calls to date).
- Published a review paper, highlighting statistical methods for uncertainty estimation in ML for property prediction.
- Developed pipeline using genetic algorithm, graph-attention ensemble to screen 30,000 small molecules for 3CL-protease binding.

PUBLICATIONS

A Benchmark Framework for Evaluating Structure-to-Sequence Models for Protein Design. Chan, J., **Chithrananda, S.**, Brookes, D. & Sinai, S., NeurIPS ML for Structural Biology Workshop (2022). Preprint to be released soon.

ChemBERTa-2: Towards Chemical Foundation Models. Ahmad, W., Simon, E., **Chithrananda, S.,** Grand, G., & Ramsundar, B., ELLIS ML for Molecules Workshop (December, 2021). *arXiv:2209.01712*

ChemBERTa: Large-Scale Self-Supervised Pre-training for Molecular Property Prediction. **Chithrananda, S.,** Grand, G., & Ramsundar, B. (NeurIPS 2020 ML for Molecules workshop). *arXiv:2010.09885*

Assigning Confidence to Molecular Property Prediction. Nigam, A., Pollice, R., Hurley, M.F., Hickman, R.J., Aldeghi, M., Yoshikawa, N., **Chithrananda, S.,** Voelz, V., & Aspuru-Guzik, A. (Expert Opinions in Drug Discovery, Taylor and Francis) 2021.

SKILLS & INTERESTS

Programming: Python, Tensorflow, Keras, PyTorch, Pandas, RDKit, sklearn, Spark. **Infra**: AWS, Google Cloud, Docker.
Talks: Delivered research talks at CMU, Baylearn, NeurIPS, Royal Society of Chemistry and Re-Work to audiences of 200+.

AWARDS, GRANTS AND FUNDING

Masason Foundation Fellow - Softbank: Support for my research and education from Masayoshi Son's foundation
Emergent Ventures Fellowship: Recipient of research fellowship, supported by the Thiel Foundation. (Press Release)
Re-Work Young Researcher: Delivered talk at conference in front of over 1000 attendees on independent research in comp. biology
Tensorflow Research Cloud Fellowship: Grant offered by Google to use TPU graphics processing pods for ML research