Seyone Chithrananda

+1 510 990 7557 | seyonec@berkeley.edu | Website | GitHub | Google Scholar | Berkeley, California

Education

University of California, Berkeley — Berkeley, California
Bachelors of Arts in Computer Science — *August 2021 – May 2025*

Major coursework: Deep Reinforcement Learning (graduate course), Efficient Algorithms and Intractable Problems, Therapeutics Discovery and Development, Intro to Machine Learning, Machine Structures, Convex Optimization, Computational Functional Genomics (graduate course), Discrete Math and Probability Theory, Data Structures, Designing Information Devices and Systems I, II (linear algebra, circuit analysis, ML), Structure and Interpretation of Computer Programs, Single/Multivariable Calculus

Work Experience

Microsoft Research — Cambridge, Massachusetts
Research Intern — *May 2023 – Present*

- Advised by Dr. Kevin Yang and Dr. Judith Amores, developing machine learning models for olfaction.
- Developed multi-modal architecture for predicting protein-ligand interactions with olfactory receptors, using geometric learning and protein language models.
- Developed model that can predict a molecules perceptual smell, using information about the olfactory receptors it activates. Preprint in progress, to be submitted to journals.

Dyno Therapeutics — Cambridge, Massachusetts
ML Design Intern — *May 2022 – August 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.
- 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 *(ML in Structural Biology, 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 — San Francisco, California (remote)
Computational Chemistry intern — *May 2021 – Aug 2021*

- Developing computational strategies for DNA-encoded library (DEL) 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.

Research Experience

Innovative Genomics Institute, Laboratory of Prof. Jennifer Doudna — Berkeley, California
Undergraduate Researcher — *January 2023 – Present*

- Student in Prof. Jennifer Doudna's laboratory working with Ph.D. student Ron Boger on various projects encompassing model-guided engineering of CRISPR-Cas enzymes and remote homology detection.
- Co-first author on protein semantic similarity search method for RNA-Guided endonuclease discovery *(Computational Biology, ICML 2023)*.

University of Toronto, Laboratory of Prof. Alan Aspuru-Guzik — Toronto, Canada
Research Student — *April 2020 – April 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 a review paper, highlighting statistical methods for uncertainty estimation in ML for property prediction. *(Expert Opinions in Drug Discovery, 2021* .
- Developed pipeline using genetic algorithm, graph-attention ensemble to screen 30,000 small molecules for SARS-COV-2 3CL-protease binding.
- Developed novel pretraining strategies for large-scale language modeling of molecules with open-source organization DeepChem, ChemBERTa. Model has 1M+ API calls, 200 citations *(ML for Molecules, NeurIPS 2020).*

Publications

R Boger*, AX Lu*, **S Chithrananda***, K Yang, P Skopintsev, B Adler, E Wallace, P Yoon, P Abbeel, J Doudna. TOPH: Adapting A Contrastive Question-Answering Framework for Protein Search. ICML Workshop on Computational Biology, 2023.

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

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

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

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

Awards & Achievements

Masason Foundation Scholar | *Website*

- Support from Masayoshi Son's foundation (founder of Softbank Group). Provided scholarships for research and tuition at UC Berkeley.

Emergent Ventures Fellow | *Press*

- Emergent Ventures is a low-overhead fellowship and grant program that supports entrepreneurs and brilliant minds with highly scalable, "zero to one" ideas for meaningfully improving society.
- Awarded two grants, one in March 2020 for research in computational chemistry, and a second to support living expenses for undergraduate studies at UC Berkeley.

Re-Work Young Researcher: | *Talk*

- Delivered talk at conference in front of over 1000 attendees on independent research in computational biology.

Scientific Talks

Rising Stars Seminar: Understanding the Combinatorial Code of Smell. Alaa Lab, UC Berkeley. Sept 1, 2023 | *Talk*

Scientific Machine Learning Webinar Series: ChemBERTa: Large-Scale Self-Supervised Pretraining for Molecular Property Prediction. Carnegie Mellon University. Feb 18, 2021 | *Talk*

Bay Area Machine Learning Symposium : ChemBERTa: Large-Scale Self-Supervised Pretraining for Molecular Property Prediction. Remote. Oct 15, 2020 | *Talk*

Re-Work Deep Learning Summit: Deep Learning to Understand Gene Expression. Montreal, Canada. Nov 3, 2019 | *Talk*

Service

Reviewing: NeurIPS Workshop on ML for Structural Biology

Co-organizer: Berkeley BioML Seminar Series | *Overview*

Machine Learning at Berkeley Supporting Berkeley undergrads to do meaningful research, through organizing research talks, reading groups, project mentorship, & organizing seminars. | *Website*

Skills

Programming: C, C++, Java, Python

Frameworks: Pytorch, Tensorflow, Pandas, Torch Geometric