Tianyu Lu

Department of Computer Science
University of Toronto
tianyu.lu@mail.utoronto.ca
https://tianyu-lu@github.io · tianyu-lu

Research Interests

I am interested in machine learning for computational biology, in particular:

- Generative models of protein sequence and structure applied to protein engineering
- Exploring $p(sequence, structure, function)$

Publications

Strokach A, **Lu T**, and Kim PM. "ELASPIC2 (EL2): Combining Contextualized Language Models and Graph Neural Networks to Predict Effects of Mutations." Journal of Molecular Biology (2021): 166810.

Conferences

Boucinha A, Kell B, Sheikh F, Diep P, Yeung A, Escobar A, Emond CA, Pierce C, Siddartha K, Chang L, Sadatmousavi P, Stephens S, **Lu T**, Sajtovich VA. "A framework towards transdisciplinary synthetic biology curricula for heterogeneous undergraduate cohorts." Canadian Engineering Education Association Conference (2021).

Preprints

Lu T, and Silva A. "dynUGENE: an R package for uncertainty-aware gene regulatory network inference, simulation, and visualization." bioRxiv (2021).

Experience

ProteinQure
Junior Machine Learning Scientist *May 2021 – Present*

- Lambda networks for computing long-range forces in linear time for noncanonical amino acids

University of Toronto
Research Student *Sep 2020 – April 2021*
Supervisor: Prof. Alan Moses

- Learning gene regulatory network dynamics using neural networks mixed with ODEs
- Designing gene regulatory networks *de novo* using automatic differentiation
- Disentangling performance of large protein language models from linear regression and embedding dimension

University of Toronto
Research Student *Aug 2019 – Sep 2020*
Supervisors: Prof. Philip Kim, Dr. Pedro Alberto Valiente Flores

- Designing novel protein folds with generative models of protein structures (Transformer, GAN).
- Designing oncoprotein inhibitors using Rosetta and molecular dynamics simulations.
- Implemented code to search the PDB for protein surfaces that mimic DNA.

iGEM Toronto
Co-President *Nov 2020 – Present*
Drylab Lead *Apr 2020 – Nov 2020*
Drylab Member *Apr 2019 – Apr 2020*
Supervisor: Prof. Radhakrishnan Mahadevan

- Contributed ML-guided designed plastic-degrading enzyme to BioBrick.
- Working on active learning methods for an iterative drylab-wetlab feedback loop
- Designed a plastic-degrading enzyme using model-based optimization.
- Quantifying benefits of learned protein sequence embeddings on protein function prediction.
- Analyzing PET catalysis dynamics with molecular dynamics simulations.

Canadian Synthetic Biology Education Research Group
Machine Learning Instructor *Sep 2019 – Present*
Supervisors: Patrick Diep, Brayden Kell

- Created code resource on machine learning for protein design and systems biology .
- Presented seven hours of content, covering both classical (Docking, Rosetta, MD) and recent methods (Sequence-to-function models, generative models, representation learning, active learning).

Education

University of Toronto
B. Sc. in Bioinformatics and Computational Biology, Computer Science *Sep 2017 – Jun 2021*
cGPA: 3.92/4.00

Skills

Programming Python, PyTorch, NumPy, Bash
Tools PyMOL, VMD, GROMACS, RosettaScripts, Unix, LaTeX

Talks

1. Accelerating Plastic Recycling with PETase, *iGEM Giant Jamboree*, Boston, MA, Oct. 2019.
2. Recurrent Neural Networks for Protein Design, *Ontario iGEM Conference*, Guelph, ON, Jul. 2019.

Awards & Grants

- Skule Endowment Fund ($5000) *2021*
- Dean's Student Initiative Fund ($1000) *2021*
- NSERC Undergraduate Research Award ($6000) *2020*
- COVID-19 Student Engagement Grant ($3000) *2020*
- iGEM Gold Medal, Best Manufacturing Project Nomination, Boston MA *2019*
- National Biology Competition, Top 1% ($3000) *2017*
- University of Toronto Entrance Scholarship ($7000) *2017*

Interests

Rachmaninoff Piano Concerto No. 2, cycling, public transport, making things with flour