Tony Shen

+1 (778) 681-9007 • github.com/tsa87 • tonyzshen@gmail.com

EDUCATION

Northeastern University / Broad Institute of MIT and Harvard — Boston, MA
Visiting researcher — Feb 2025 - Present
- **Research:** Foundation model for bio-molecular generation, and antibiotics discovery applications.
- **Supervisor:** Dr. Wengong Jin

Simon Fraser University — Burnaby, BC
PhD in Computing Science — May 2023 - Present
- **Research:** Structure-based drug design with generative flow networks and flow matching models.
- **Supervisor:** Dr. Martin Ester

BSc Joint Major in Computing Science & Molecular Biology and Biochemistry (GPA: 3.97/4.33) — Jan 2019 - May 2023

SELECTED PUBLICATIONS

- **Tony Shen***, Seonghwan Seo*, Ross Irwin, Kieran Didi, Simon Olsson, Woo Youn Kim, Martin Ester. Compositional Flows for 3D Molecule and Synthesis Pathway Co-design. **Under conference submission, ICLR GEM and AI4MAT workshop**, 2025.
- Seonghwan Seo, Minsu Kim, **Tony Shen**, Martin Ester, Jinkyoo Park, Sungsoo Ahn, Woo Youn Kim. Generative Flows on Synthetic Pathway for Drug Design. **International Conference on Learning Representations (ICLR)**, 2025.
- **Tony Shen**, Seonghwan Seo, Grayson Lee, Mohit Pandey, Jason R Smith, Artem Cherkasov, Woo Youn Kim, Martin Ester. *TacoGFN*: Target Conditioned GFlowNet for Structure-Based Drug Design. **Transactions on Machine Learning Research (TMLR)**, 2024.

WORK EXPERIENCES

Transcripta Bio — Palo Alto, CA
Machine Learning Consultant (Part-Time) — Sep 2023 - Dec 2024
- Leading experiments on generative models for molecules with desired gene expression, resulting in the discovery of several novel molecules with better predicted bioactivity profile.
- Enabling company's first ML-based virtual screening pipeline with compound library standardization, characterization and featurization - which resulted in several wet-lab validated hits.

Recursion — Salt Lake City, UT
Software Engineer (Part-Time) — Sep - Dec 2022
Software Engineering Intern — Jun - Sep 2022
- Implemented gene search and enrichment analysis of biological pathways with **React** and Python using known biology databases to accelerate drug-target discovery workflow. Setup SLO alerts using **OpenTelemetry** and **Prometheus**.
- Built full-stack graph visualization and analysis app for millions of relationships between biological entities using **Dash**.

Amgen — Burnaby, BC
Bioinformatics Co-op — Jan - Apr 2022
- Developed binding affinity prediction model of antibody and target protein on cell sorting data using **PyTorch** and ESM.
- Built Dash web app to visualize immune repertoire containing thousands of sequences through network graph.

Google — Waterloo, ON
Software Developer Intern — May - Aug 2021
- Introduced contrastive learning to Retail AI ranking model and reduced average first click position by 0.36 using **TensorFlow**.
- Implemented adaptive multi-objective loss balancing algorithms (GradNorm, Uncertainty Weighting) to reduce over-fitting in multi-task ranking model.
- Used **C++** and **Flume (Apache Beam)** to build metric aggregation pipelines that processed terabytes of search logs to provide insights on search quality.

Georgian — Toronto, ON
Machine Learning Engineering Intern — Sep - Dec 2020
- Built **ML Platform** command line tool that enables data scientists to seamlessly train hundreds of **docker**ized reproducible experiments on cloud in one command.
- Trained and evaluated models predicting startups' raise likelihood to help investment team on candidates prioritization.
- Generated novel features from large news dataset using **Apache Spark** and **increased AUPR by 10%** compared to baseline.

Google — Waterloo, ON
STEP Intern — May - Aug 2020
- Proposed and implemented a full stack web & mobile platform that enables non-technical users to create functional bots.

- Owned and developed **Flutter/Dart** application and **Cloud Firestore** component. Proposed and communicated key technical designs in schema and framework choice, resulting in accelerated completion of cross-platform app.
- 20% Project: Worked with Cloud Health team to generate human mutation histogram, using **SQL** and **Python** to pull data from Genome Aggregation Database in **BigQuery**.

TECHNICAL SKILLS

Programming Languages: Python, C/C++, Java, JavaScript, SQL.

Technical Tools: PyTorch, Tensorflow, Spark, Beam, Flutter, React, Dash, PyMol, RDKit, Docker, Google Cloud, AWS.

Research Interest: AI for Drug Discovery, Generative Models, Flow Matching, Diffusion Models, Geometric Deep Learning, GFlowNets.

COMMUNICATIONS

Guest interviewer - InterBios Podcast May 2024
Interviewed Dr. Anthony Fejes - CEO of HTuO Biosciences on the role of molecular simulation on drug discovery.

Guest writer - Valence Portal Blogs Apr 2024
Published article on AI for Drug Design and Generative Flow Networks (link)

SUPERVISING

Grayson Lee Oct 2023 - May 2024
Geometric-informed GFlowNets for Structure-Based Drug Design - **MoML 2024 (Spotlight)**

ACADEMIC SERVICE

Reviewer
Transactions on Machine Learning Research (TMLR) Oct 2024 -

PROJECT EXPERIENCES

SFU Community Trust
Summer Intern / Project Founder May - Aug 2019
- Created a pilot for co-living residence by converting existing units and provided affordable housing to 8 students.
- Presented feasibility report to developers and Campus Planning Director, resulting in elevated interest for coliving concepts.

InHouse Mobile
Founder Jan - May 2020
- Developed mobile ordering platform for restaurants during COVID-19 - using **React Native/Java Script**, **Expo** and **MongoDB**, and delivered both customer facing and restaurant-end apps with-in 3 months.
- Led a 4-student team and improved development process through weekly code sprints, code reviews and bi-weekly standups.

AWARDS

Undergraduate Open Scholarship 2020 - 2023
Semester GPA above 4.0 (equivalent to A) in 2020 Spring, 2020 Fall, 2021 Spring, 2021 Fall, 2022 Fall and 2023 Spring.

HackSeqRNA 2020 - 1st Place May 2020
- Created, trained and evaluated 1-D Convolutional models on RNA targets to predict ASO effectiveness with **Keras**.

SFU Waste Sorting Robot Competition - 2nd Place Mar 2019

VANTEC Angle Investor Network Pitch Night - 2nd place Nov 2018
- Pitched coliving student housing project and won 2nd place out of 7 upcoming startups.