Kyle Roh with background in machine learning and quantitative biological systems, interested in applying this skillset to quantitative research roles in finance.

929-319-7120 rohk@stanford.edu linkedin.com/in/kun-hyung-r

Education

Stanford University **Expected: 2027**

B.S in Computer Science, minor in Statistics *Stanford, CA*

Relevant Coursework (*: Graduate Coursework)

- Deep Learning*
- Computational structural biology*
- Probability Theory
- NLP w/ Deep Learning*
- Organic Chemistry
- Biophysics*
- C++ algorithms, systems
- PDE, ODE, Abstract

Linear, Multivar Calc, Convex Optimization

Experience

Mt. Sinai School of Medicine **Oct 2020 – Sep 2024**

Paid full-time researcher *New York, NY*

- Developed 91% accuracy expression-based ML drug discovery platform (Pytorch) to discover small molecules that modulate expression of specific genes. Discovered hits were validated in lab. Overcame low-data availability using Borderline SMOTE and Directed MPNNs.
- Discovered first-in-class candidate drugs against Alzheimer's Disease pathology and age-related diseases using machine learning. The experimental results outperformed all previously reported protective drugs and in clinical trials.
- Led chemogenomic screening of 140,000 compounds in vivo by training team of 10 new researchers for target identification.
- Submitted to Nature, published 6 other peer-reviewed international conferences. Pending 2 publications.

Talo Labs **Aug 2022 – Dec 2022**

ML Engineering Intern *Singapore*

- Engineered machine learning pipeline using Pytorch predicting QM-based force field interactions. (Received data from molecular dynamics simulations)
- Increased biophysical drug discovery platform accuracy by over 20%.
- Collaborated with team members to generate compounds targeting JAK2 and Kir 4.1.

Stanford Mathematics **2023**

Math Researcher *Palo Alto, CA*

- Research on Jordan Curves and Abstract Linear Algebra under Dr. Rubinstein-Salzedo. Variations of the Toeplitz conjecture using homeomorphisms of Jordan curve parameterization in R^n and Sobolaev spaces.
- Received lectures on p-adic analysis, Algebraic Topoology, and Probability Theory

Projects

D-MPNNs for probabilistic classification | *Python, Pytorch*

∗ Performed EDA to analytically reduce high dimensional dataset containing genetic signature and respective chemicals.
∗ Used random forest, logistic regression, GNN, SMOTE for classification of imbalanced data.
∗ Used reverse-docking to identify unknown physical target of a newly discovered compound. Then validated binding using free energy calculations done by DockThor, powered by supercomputer SDumont.

Self-reinforcing LLMs for de novo drug generation | *Python, Pytorch*

∗ Incorporated pairwise cringe loss (PCO) to reinforce modified LLaMa v3.1 for small molecule generation.
∗ Outperformed baseline models in molecule novelty and validity, improvement in inhibition properties

Technical Skills

Languages: Python, C++, VBA, R
Machine Learning: Pytorch, TensorFlow, Keras, Numpy, Pandas, Sklearn
Bioinformatics : HTVS, Transcriptomics, AutoDock, VMD, GROMACS, CHARMM

Awards / Honors

∗ **Research Awards**: 1st Place Winner, 2x Finalist ISEF;
Top 40 Finalist, Science Talent Search (Formerly known as Westinghouse or Intel)
∗ **Math**: 2x AIME Qualifier, 5th Place Team ARML, 15th Place Team HMMT (Harvard-MIT Math Tournament), Individual top scorer.
∗ **Science**: USA Biology Olympiad Top 100