

Yusuf H. Roohani

Website: cs.stanford.edu/~yhr

Email: yroohani@stanford.edu

EDUCATION	<p>Stanford University, Stanford, CA Jan 2020 - Present Ph.D., Biomedical Data Science GPA: 4.10/4.0 Dissertation Title: <i>Cell Engineering: From Differentiable Biology to Biological Design</i> Advisors: Jure Leskovec, Stephen Quake Committee: Barbara Engelhardt, Caroline Uhler, Daphne Koller</p> <p>Carnegie Mellon University, Pittsburgh, PA Jan 2014 - Aug 2015 M.S., Mechanical Engineering. Advisors: Peter Adams, Allen Robinson GPA: 4.0/4.0 Vellore Institute of Technology, Vellore, India Jul 2009 - Jun 2013 B.Tech., Mechanical Engineering GPA: 8.81/10</p>
HONORS AND AWARDS	<p>Poster Award: Top 3 out of 200 at Stanford BioX Interdisciplinary Initiatives (\$500) (2023) Innovation Award: Society for Lab Automation and Screening (\$10K) (2023) Best Poster: (Machine Learning Track) Intelligent Systems for Mol. Biology (ISMB) (2022) Best Poster: Single Cell Genomics meets Data Science, Munich (\$500) (2022) Full PhD funding awarded by GSK, including tuition + full-time pay (\$1M+) (2019) GSK Exceptional Science Award Deep learning for cellular images (\$17K) (2018) Undergraduate Research Assistantship Birck Nanotechnology Centre, Purdue (2013) Merit Certificates, Academic Excellence (x4), Vellore Inst. of Tech. (2010/11/12/13)</p>
JOURNAL PUBLICATIONS	<p>[Nature Biotechnology] [Code] GEARs: Predicting transcriptional outcomes of novel multi-gene perturbations, 2023. Roohani, Y., Huang, K., Leskovec J.,</p> <p>[Nature Methods] Towards Universal Cell Embeddings: Integrating Single-cell RNA-seq Datasets across Species with SATURN, 2023. Rosen Y.*, Brbić M.*, Roohani, Y.*, Swanson K., Li Z., Leskovec, J..</p> <p>[Nature Chemical Biology] Artificial Intelligence Foundation for Therapeutic Science, 2022 Huang, K., Fu, T., Gao, W., Zhao, Y., Roohani, Y., Leskovec, J., ... Zitnik, M..</p> <p>[J Ultrasound Med.] Enhanced point-of-care ultrasound applications by integrating automated feature-learning systems using deep learning, 2018. Shokoochi H., LeSaux M., Roohani Y., Litepio A., Huang C., Blaiwas M.</p> <p>[Atmospheric Environment] Impact of natural gas development in the Marcellus and Utica Shales on regional ozone and fine particulate matter levels, 2017. Roohani, Y., Roy, A., Heo, J., Robinson, A., & Adams, P.</p>
CONFERENCES	<p>[NeurIPS] Zero-shot causal learning, 2023. Nilforoshan H.*, Moor M.*, Roohani Y., Chen Y., Šurina A., Yasunaga M., ... Leskovec J..</p>

[[NeurIPS](#)] Therapeutics Data Commons: Machine Learning Datasets For Therapeutics, 2021. Huang, K., Fu, T., Gao, W., Zhao, Y., **Roohani, Y.**, Leskovec, J., ... Zitnik, M..

[[NeurIPS ML4H Workshop](#)] Predicting Language Recovery after Stroke with Convolutional Networks on Stitched MRI, 2018.

Roohani Y., Sajid N., Hope T., Price C., Madhyastha P.,

PREPRINTS

[[biorXiv](#)] Universal Cell Embeddings: A Foundation Model for Cell Biology, 2023.

Rosen, Y.*, **Roohani, Y.***, Agarwal, A., Samotorcan L., Quake, S., Leskovec J.

[[arXiv](#)] CausalBench: A Large-Scale Benchmark for Network Inference from Single-Cell Perturbation Data, 2022.

Chevalley, M., **Roohani, Y.**, Mehrjou, A., Leskovec, J., Schwab, P.

[[arXiv](#)] On the opportunities and risks of foundation models, 2021.

Bommasani, R., Hudson, D. A., ... **Roohani, Y.**, ... Liang, P.

INVITED
TALKS

Cell Engineering: From differentiable biology to biological design

AstraZeneca, Cambridge, UK (2024)

Arc Institute, Palo Alto, CA (2023)

EvolutionaryScale, New York, NY (2023)

Universal cell embeddings for predicting multigene perturbation outcomes

Stanford Graph Learning Workshop, Department of Computer Science, Stanford, CA (2023)

GEARS: Predicting transcriptional outcomes of novel multi-gene perturbations

Genentech Research and Early Development, South San Francisco, CA (2023)

Altos Labs, Palo Alto, CA (2022)

Stanford Network Biology Journal Club (Kundaje/Bassik Lab, Stanford Genetics/CS) (2022)

Recursion Pharmaceuticals, Salt Lake City, UT (2022)

Predicting outcomes of multi-gene perturbations and identifying optimal inter-

ventions Stanley Qi Lab Group Meeting (Stanford Bioengineering) (2022)

Assessing biological diversity of a compound collection using high throughput

cellular imaging, Society for Lab Automation and Screening Conference. (2020)

Data Analytics and Machine Learning in Drug Discovery.

Guest Lecture, School of Engineering, Tufts University. (2019)

Accelerating High Throughput Drug Discovery Using Deep Learning.

ReWork, Deep Learning for Healthcare, Boston. (2018)

CONTRIBUTED
TALKS

GEARS: Predicting transcriptional outcomes of novel multi-gene perturbations

Society for Lab Automation and Screening Conference [**Innovation Award**] (2023)

Machine Learning for Computational Biology [**17% Acceptance Rate**] (2022)

Single Cell Genomics meets Data Science, Munich [**Best Poster Award**] (2022)

CRISPR Perturbations and Beyond, Wellcome-Sanger Institute (2022)

WORK EXPERIENCE	<p>Stanford University, Stanford, CA PhD Student Jan 2020 - Present Predicting outcomes of genetic perturbations and identifying optimal interventions. Employed at GSK during PhD at Stanford. No restrictions on PhD research. See Awards</p> <p>GlaxoSmithKline, Cambridge, MA Manager, Machine Learning Engineer (Level 6) Apr 2019 - Jan 2020 Investigator (Level 7) Nov 2017 - Mar 2019 Data Scientist (Level 8) Jul 2016 - Oct 2017 Led a cross-disciplinary team on a project to biologically profile their 2M+ compound collection using complex multi-modal datasets and high throughput screening. Applied to an active program for discovering new biological targets.</p> <p>Merrimack Pharmaceuticals, Cambridge, MA Computational Modeler Intern Sep 2015 - Apr 2016 Developed dynamic system models to mechanistically simulate signaling networks in cancer. Compared results against patient data to identify biomarkers for patient stratification.</p> <p>Tata Industries, Mumbai, India Technical Analyst Intern Sep 2013 - Nov 2013 Advised the strategic venture capital division on investments in materials science research</p>																												
REVIEWING	PLOS Computational Biology, NeurIPS, NeurIPS: AI for Science, MICCAI.																												
GRANTS RECEIVED	California Institute for Regenerative Medicine (Co-recipient), \$700K 2021 - 2023 UCB-Stanford Seed Funding Program, \$50K 2023 - 2024																												
SELECTED COURSES (STANFORD)	<table border="0"> <thead> <tr> <th>Course</th> <th>Number</th> <th>Grade</th> <th>Term</th> </tr> </thead> <tbody> <tr> <td>Machine Learning under Distribution Shift</td> <td>CS 329D</td> <td>A+</td> <td>Spring 2021</td> </tr> <tr> <td>Design and Analysis of Algorithms</td> <td>CS 161</td> <td>A+</td> <td>Summer 2020</td> </tr> <tr> <td>Machine Learning</td> <td>CS 229</td> <td>A</td> <td>Spring 2020</td> </tr> <tr> <td>Fundamentals of Real Analysis</td> <td>MATH 171</td> <td>A</td> <td>Spring 2020</td> </tr> <tr> <td>Artificial Intelligence</td> <td>CS 221</td> <td>A</td> <td>Fall 2020</td> </tr> <tr> <td>Machine Learning with Graphs</td> <td>CS 224W</td> <td>A</td> <td>Fall 2019</td> </tr> </tbody> </table>	Course	Number	Grade	Term	Machine Learning under Distribution Shift	CS 329D	A+	Spring 2021	Design and Analysis of Algorithms	CS 161	A+	Summer 2020	Machine Learning	CS 229	A	Spring 2020	Fundamentals of Real Analysis	MATH 171	A	Spring 2020	Artificial Intelligence	CS 221	A	Fall 2020	Machine Learning with Graphs	CS 224W	A	Fall 2019
Course	Number	Grade	Term																										
Machine Learning under Distribution Shift	CS 329D	A+	Spring 2021																										
Design and Analysis of Algorithms	CS 161	A+	Summer 2020																										
Machine Learning	CS 229	A	Spring 2020																										
Fundamentals of Real Analysis	MATH 171	A	Spring 2020																										
Artificial Intelligence	CS 221	A	Fall 2020																										
Machine Learning with Graphs	CS 224W	A	Fall 2019																										