Hossein Sharifi-Noghabi

Machine Learning Researcher — Biomedical Researcher
hsharifi@sfu.ca | Homepage | LinkedIn | Github

Highlight

- Expertise in developing and applying machine learning methods for real-world problems (six years of experience in employing machine learning in biology and medicine with five top-tier publications).
- Experienced in teamwork particularly when members have diverse backgrounds (four years of experience in collaborations with life scientists in academia and one year of cross-team collaboration in industry).
- Interested and passionate about learning (More than 12 online certificates on different topics such as deep learning and reinforcement learning).
- Skilled in Python, R, Pytorch, Slurm, Bash, scikit-learn, Git, Matplotlib, and Pandas (five years of experience in deploying them for real-world problems).

Education

Simon Fraser University — Burnaby, Canada
Ph.D. in Computer Science — *Sep. 2016 – Sep. 2021*

Ferdowsi University of Mashhad — Mashhad, Iran
M.Sc. in Artificial Intelligence — *Sep. 2012 – Feb. 2015*

Sadjad University of Technology — Mashhad, Iran
B.Eng. in Information Technology — *Sep. 2008 – Jul. 2012*

Experience

AI4Life Resident (Internship) — Sep. 2021 – Sep 2022
Novartis — *Basel, Switzerland*

- Design and development of a model for batch effect correction across blood-based proteomics data.
- Implemented codes in Python, PySpark, and utilized Pytorch and Matplotlib.

Research Student (Co-op) — June 2020 – April 2021
Princess Margaret Cancer Centre — *Toronto, Canada*

- Led a project to propose guidelines on how to employ machine learning in pharmacogenomics. These guidelines investigate generalization of machine learning methods in within-/cross-domain drug response prediction using RNA-Seq data and provides solutions to improve that.
- Implemented codes in Python and R, and utilized Pytorch and scikit-learn.

Research Assistant — Sep. 2016 – July 2021
Simon Fraser University — *Burnaby, Canada*

- Led multiple projects on improving the accuracy of drug response prediction in patients using gene expression data. I developed 3 methods based on multi-modal representation learning, transfer learning with input and output space adaptation, and semi-supervised domain generalization. On average, these methods improved the prediction accuracy by 9% compared to state-of-the-art methods.
- Implemented codes in Python and utilized Pytorch and Keras frameworks.
- Mentored undergraduate research assistants in Database and Data Mining Laboratory.
- Collaborated with life scientists, clinical fellows, and staff at the Vancouver Prostate Centre.

Collaboration — Sep. 2017 – Jan. 2018
Decipher Biosciences, Inc — *Vancouver, Canada*

- Developed a novel method to predict metastasis (a binary outcome) in prostate cancer from early stages using gene expression data. This method was based on denoising autoencoders and transfer learning on unlabelled and labelled clinical samples and improved the prediction accuracy by 2% compared to state-of-the-art clinical studies.
- Implemented codes in Python and R, and utilized Tensorflow framework.

Selected Publication

H. Sharifi-Noghabi, P. Alamzadeh Harjandi, O. Zolotareva, C. Collins, M. Ester, **Out-of-distribution generalization from labelled and unlabelled gene expression data for drug response prediction** *Nature Machine Intelligence 3, 962–972 2021.*

H. Sharifi-Noghabi, S. Jahangiri-Tazehkand, P. Smirnov, C. Hon, A. Mammoliti, S. Kadambat Nair, A. Mer, M. Ester, B. Haibe-Kains, **Drug Sensitivity Prediction From Cell Line-Based Pharmacogenomics Data: Guidelines for Developing Machine Learning Models** *Briefings in Bioinformatics 22(6), 1–14 2021.*

O. Snow, H. Sharifi-Noghabi, J. Lu, O. Zolotareva, M. Lee, M. Ester, **Interpretable drug response prediction using a knowledge-based neural network** *Proceedings of the 27th ACM SIGKDD 2021.*

H. Sharifi-Noghabi, S. Peng, O. Zolotareva, C. Collins, M. Ester, **AITL: Adversarial Inductive Transfer Learning with input and output space adaptation for pharmacogenomics** *Bioinformatics 36, i380-i388 (ISMB 2020).*

H. Sharifi-Noghabi, O. Zolotareva, C. Collins, M. Ester, **MOLI: multi-omics late integration with deep neural networks for drug response prediction** *Bioinformatics 35 (14), i501–i509 (ISMB/ECCB 2019).*

H. Sharifi-Noghabi, H. Rajabi Mashhadi, and K. Shojaee, **A novel mutation operator based on the union of fitness and design spaces information for Differential Evolution**, *Soft Computing (21) 6555–6562 2016.*

M. Mohammadi, H. Sharifi-Noghabi, H. Rajabi Mashhadi, and G. Hodtani, **Robust and stable gene selection via Maximum-Minimum Correntropy Criterion** *Genomics (170) 83-87 2016.*

For the complete list please visit my Google Scholar

Skill Set

- **Languages**: Python, R, MATLAB
- **Frameworks**: Tensorflow, Keras, Pytorch, PySpark
- **Machine learning**: Transfer learning, Domain generalization and adaptation, Multi-task learning, Adversarial learning, Semi-supervised learning, Meta-learning, Multi-modal data integration
- **Libraries**: Pandas, Numpy, Scikit-learn, Matplotlib
- **Other**: High Performance Computing (Slurm), Git, Bash, Object-oriented programming in Python, Teamwork (Slack, MS Teams, Miro), Visualization (Photoshop, Lucidchart), and Latex

Research Interests

- **Transfer Learning**: Out-of-distribution generalization, Domain adaptation, Multi-task learning
- **Learning with limited data**: Semi-supervised learning, Meta-learning, Self-supervision
- **Bioinformatics**: Pharmacogenomics, Precision medicine, Multi-omics integration
- **Other**: Federated learning (for personalization and under data/system heterogeneity), Reinforcement learning (sequential decision making under distribution shifts)

Selected Projects

- **Drug Discovery**: Automatic chemical design via Variational AutoEncoders (VAE) using SMILES representation.
- **Feature Selection**: Biomarker discovery and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins via ensemble feature selection and classification.
- **Vision**: Applied Convolutional Neural Networks to image verification and recognition tasks and neural style transfer to generate new arts.
- **NLP**: Applied LSTM and GRU models to synthesize Shakespeare's text, speech recognition, and music synthesis.
- **RL**: Implemented Q-Learning and Expected Sarsa both with ϵ-greedy action selection on Cliff World.
- **RL**: Implemented semi-gradient TD with a Neural Network as the function approximator for an RL agent in 500-State Random Walk Environment.