Rahul Mehta

✆ +1 630 936 9567
✉ rmehta1987@gmail.com
🖥 rmehta1987.github.io/

Education

2013-2020	**Ph.D in Bioengineering**, *University of Illinois at Chicago*, Chicago,IL, Dissertation: Patterns Among and Between Somatic Mutations and Medical Imaging in Human Cancers.
2005–2011	**Bs.C in Electrical Engineering**, *University of Illinois at Urbana Champaign*, Urbana-Champaign,IL.

Skills

Computing	Python, BASH, Git, AWS (S3, EC2, and RDS), SLURM, C++, Latex, PostgresSQL, Docker
Machine Learning API and Methods	PyTorch, openCV, Likelihood-free inference, Diffusion models, Computer Vision, Self-supervised learning, Recurrent Neural Networks, Transformers (LLM)
Bioinformatics API	SAMtools, PLINK, Hidden-Markov Models, SnakeMake, RFMix
Analysis	Multimodal datasets, latent-space/unsupervised analysis, genetic admixture, genetic risk scores, parallel computing, time series
Medical Imaging	Magnetic Resonance Imaging, Computed Tomography, Positron Emitted Tomography, Fluorescence in situ hybridization, DICOM, PACS

Experience

2023-   **Genomic Data Science Fellow**, *Ancestry*.
- Developed a multimodal deep learning pipeline that integrates genomic data from more than 2 million people (800,000 genotyped markers) and 100 million family trees to reveal recent, fine-scale population histories of communities that share genetic ancestry.
- Customers of Ancestry will get to visualize possible origin of their communities, genetic proportions contributed by their founding ancestors, and number of genealogical founders.

2020-2023   **Postdoctoral Scholar**, *Department of Human Genetics*, University of Chicago.
- Created a memory efficient data loading technique for whole genome sequencing data from the UK Biobank GWAS consisting of 150 million alleles that included very rare allele frequencies ($> .0001$)
- Developed a novel deep learning generative model that incorporates whole genome LD blocks and GWAS to analyze how natural selection impacts complex traits
- Method shows that the distribution of natural selection is unique across traits.

2013-2020 **Graduate Research Assistant**, *Center for Magnetic Resonance Research*, University of Illinois at Chicago.
- Showed diffusion weighted MRI images offers richer information about the heterogeneity of a lesion for predicting malignancy of a breast lesion by developing a machine learning model to extract features from diffusion weighted MRI images.
- Generated accurate predictions of radiotherapy treatment response in cholangiocarcinoma using a machine learning model and 3D computer vision algorithm to extract 3D features of PET-CT medical images.
- Predicted somatic mutations from only medical images using the TCGA dataset (20 cancers, 10000 mutations) by designing a multi-modal deep generative model for the simultaneous analysis of medical images and somatic mutation data.
- Designed an innovative probabilistic model with efficient Bayesian inference that improved the analysis of co-occurring or mutually exclusive cancerous somatic mutations.

Papers

Journal and Conference Papers

2023 Mehta, R., M. Karaman, Y. Bu, Z. Zhong, W. Shiwei, C. Zhou, H. Weihong, X. Maosheng, and Z. J. Xiaohong. "Characterization of breast lesions using multi-parametric diffusion MRI and machine learning." *In Physics in Medicine and Biology* (2023).

2020 Mehta, R. and M. Karaman. "Correlated Mixed Membership Modeling of Somatic Mutation Profiles". *In International Joint Conference on Neural Networks*. IEEE. 2020.

Mehta, R., M. Karaman, and Y. Lu. "Mapping of Lesion Images to Somatic Mutations". *In Data and Text Mining in Biomedical Informatics*. ACM. 2020.

2017 Mehta, R., K. Cai, N. Kumar, M. G. Knuttinen, T. M. Anderson, H. Lu, and Y. Lu. "A lesion-based response prediction model using pretherapy PET/CT image features for Y90 radioembolization to hepatic malignancies". *In Technology in cancer research & treatment* 16.5 (2017), pp. 620–629.

Workshop Papers

2018 Mehta, R. and H. Lu. "Bayesian Power Law Models for Somatic Mutation Profiles". *In the ISMB workshop on Machine Learning in Computational and Systems Biology*. 2018.

Mehta, R. and H. Lu. "Normalized Random Measure Mixture Models in Variational Autoencoders". *In the NeurIPS workshop on Advances in Approximate Bayesian Inference*. 2018. URL: http://approximateinference.org/2018/accepted/MehtaLu2018.pdf.

Mehta, R. and H. Lu. "Power Law Models in Somatic Mutation Profiles". *In the IJCAI workshop on Biomedical Informatics with Optimization and Machine Learning*. 2018.

Manuscripts in Progress

Mehta, R. and J. Berg J. *Deep Generative Modeling of Pleiotropy and Evolutionary Parameters using Genome Wide Association Summary Statistics*.

Mehta, R., Y. Wang, and K. Noto. *Inference of (very) Recent Admixture*.

Presentations

2022 **Deep Generative Modeling Pleiotropy and Natural Selection in Complex Traits.**
Probabilistic Modeling in Genomics, 2022. **Poster**

2020 **Correlated Mixed Membership Modeling of Somatic Mutations.**
International Joint Conference on Neural Networks, 2020. **Contributed Talk**

2020 **Diffusion-Weighted MRI-Based Quantitative Markers for Characterizing Breast Cancer Lesions Using Machine Learning.**
The International Society for Magnetic Resonance in Medicine, 2020. **Poster**

2019 **Non-parametric Models of Somatic Mutation Profiles.**
IEEE Biomedical Health Informatics Special Session on Nonparametric Statistics in Omics Applications, 2019. **Invited Talk**

2018 **Power Law Models in Somatic Mutation Profiles.**
IJCAI workshop on Biomedical Informatics with Optimization and Machine Learning, 2018. **Contributed Talk**

2018 **Bayesian Power Law Models for Somatic Mutation Profiles.**
Intelligent Systems for Molecular Biology, 2018. **Poster**

2018 **Normalized Random Measure Mixture Models in Variational Autoencoders.**
NeurIPS workshop on Advances in Approximate Bayesian Inference, 2018. **Poster**

2014 **Computer Aided Response Prediction Based on Pre-therapy FDG PET/CT Imaging Biomarkers of Y90-SIRT Therapy in Patients with Primary and Metastatic Liver Cancers.**
Radiological Society of North America, 2014. **Contributed Talk**