# Gautam “Machi” Machiraju

Oakland, CA
614-599-6288
gmachi@stanford.edu

Motivated by scientific and human-centered applications, my research interests are in (1) the interpretability of foundation models to understand their reasoning processes and decision-making, (2) interpretable AI deployments in discovery settings, and (3) controlled generation using AI-identified domain knowledge. Please visit my website for more details: gmachiraju.github.io.

## EDUCATION

2018 – 2024 | **Stanford University** — Ph.D. Biomedical Informatics

*Department:* Biomedical Data Science (BDS)
Advisors: Parag Mallick (Radiology), Christopher Ré (CS) | Committee: Sylvia Plevritis (BDS, Radiology), James Zou (BDS)
*Thesis:* Going beyond salience: deep learning for the discovery of meaningful pathomic markers

2012 – 2016 | **University of California, Berkeley** — B.A. Applied Mathematics

*Concentration:* Mathematical Biology, with coursework emphasis in numerical methods, optimization, & algorithms
*Minor:* Bioengineering

## RECENT WORK EXPERIENCE

Oct 2024 – Present | **Stanford AI Lab** — Postdoctoral Scholar

• Interpretability & knowledge discovery for protein language models (PLMs)
• Knowledge discovery for remote sensing foundation models to combat human trafficking in the Amazon Rainforest
• Controlled generation for LLM-based mathematical reasoning

Aug 2018 – Sep 2024 | **Stanford AI Lab** — Graduate Student Researcher

• Built AI systems that efficiently provide evidence and rationale for their decision-making; deployments in various high-dimensional data modalities: text documents, pathology images, protein sequences & structures, and spatial biology
• Moonlighting work in global health: projects in predicting maternal/child health outcomes with satellite imagery, mobile-based mpox screening, embedding models for clinical medicine

Summer – Fall 2023 | **IBM Research** — Student Researcher

Collaboration with Stanford AIMI Center to enable visually grounded foundation models; worked on open vocabulary object detection, referring image segmentation, and multimodal fact-checking

Sep 2016 – Aug 2018 | **Stanford Radiology** — Bioinformatics Research Assistant

Projects in: physics-based models for cancer systems biology, NLP for mining biomedical literature, and anomaly detection for multi-omic time series

Summer 2016 | **Strateos** — Bioengineering SDE Intern

Software engineering and scientific computing to refine robotic platform for wet-lab automation; projects in numerical optimization, queueing, and search

## RELEVANT SKILLS

- Programming: Python (Pytorch), shell, cluster computing (Slurm), cloud computing (GCP), W&B
- Design: LaTex, vector graphics (BioRender, Adobe)
- Mentorship: Mentored 15 URM trainees on independent research projects

## PUBLICATIONS & PREPRINTS

### Transparent AI Systems for High-dimensional Biomedical Data (thesis work)

1. Gautam Machiraju, James Zou, Christopher Ré, Parag Mallick. Toward Evidence-based & Rational AI Systems for Biomedicine. *In preparation.*
2. Gautam Machiraju, Sylvia Plevritis, Fiona Ginty, James Zou, Christopher Ré, Parag Mallick. Identifying Pathomic Biomarkers of Lung Cancer Aggressiveness with Geometric AI. *In preparation.*
3. Gautam Machiraju, Alexander Derry, Arjun Desai, Neel Guha, Amir-Hossein Karimi, James Zou, Russ Altman, Christopher Ré, Parag Mallick. Prospector Heads: Generalized Feature Attribution for Large Models & Data. *International Conference on Machine Learning (ICML), 2024.*
4. Gautam Machiraju, Arjun Desai, James Zou, Christopher Ré, Parag Mallick. Prospectors: Leveraging Short Contexts to Mine Salient Objects in High-dimensional Imagery. *International Conference on Machine Learning (ICML) workshop on Interpretable Machine Learning for Healthcare (IMLH), 2023.*
5. Gautam Machiraju, Sylvia Plevritis, Parag Mallick. A Dataset Generation Framework for Evaluating Megapixel Image Classifiers & their Explanations. *European Conference on Computer Vision (ECCV), 2022.*

### Training & Fine-tuning Foundation Models for Biomedicine

6. Alexander Thieme, Tahir Miriyev, Gautam Machiraju, et al. A Medical Vision-Language Model Trained on a Large Dataset of Open Access Medical Images. *In Review.*
7. Varun Tandon, Gautam Machiraju, Parag Mallick. Grammar Matters: Exploring Grammatical Variation's Role in Improving Fine-Tuned LLMs for Biomedical Relation Extraction. *In Review.*

### ML for Public Health & Clinical Decision Support

8. Alexander Thieme, Yuanning Zheng, Gautam Machiraju, et al. Development and Evaluation of an Image-based Deep Learning Algorithm to Classify Skin Lesions from Mpox Virus Infection. *Nature Medicine, 2023.*
9. Minh Nguyen, Conor Corbin, Tiffany Eulalio, Nicolai Ostberg, Gautam Machiraju, Ben Marafino, Michael Baiocchi, Christian Rose, Jonathan Chen. Developing Machine Learning Models to Personalize Care Levels among Emergency Room Patients for Hospital Admission. *Journal of the American Medical Informatics Association (JAMIA), 2021.*

### Mathematical & Statistical Modeling for Cancer Systems Biology

10. Hunter Boyce, Gautam Machiraju, Parag Mallick. Spatial Statistics for Spatial Biology. *In Review.*
11. Gautam Machiraju, Parag Mallick, Hermann Frieboes. Multicompartment Modeling of Protein Shedding Kinetics During Vascularized Tumor Growth. *Nature Scientific Reports, 2020.*

## AWARDS, SCHOLARSHIPS & FELLOWSHIPS

- Creativity in Research Scholars (CIRS) Program, Stanford d.school (2023 cohort)
- Canary-ACED Graduate Fellowship (100% support) via International Alliance for Cancer Early Detection (2022-2023)
- Stanford Data Science Scholarship (50% support) via Stanford Data Science Institute (2021-2022, 2023-2024)
- 2022 HAI Google Cloud Credits Award ($40K) for "Algorithmically identifying histo-molecular biomarkers of cancer progression" (2022-2023)
- National Library of Medicine (NLM) Training Grant (100% support) via the NIH (2018-2022)

## ACADEMIC TEACHING, SERVICE, & LEADERSHIP

- **Research Mentor**, SMASH Rising. Co-mentoring four undergraduate students of underrepresented backgrounds

in STEM. Taught programming and proteomics background material; led research to identify proteins shed into blood with in-house mouse xenografts. June 2022 – August 2022.

- Reviewer, European Conference on Computer Vision (ECCV), 2022.
- Reviewer, Annual Biomedical Research Conference for Minority Students (ABRCMS), 2022.
- Mentor, Stanford ADVANCE Mentor Collective Program (in partnership with Mentor Collective). Advising mentees from around the US who are applying to colleges or graduate schools. Part of the Stanford ADVANCE volunteer group. Received training and certification to be on-call as a mentor when requested. June 2021 – Present.
- Co-Czar, Biomedical Informatics Training Program. Elected student leader and liaison working on recruitment, admissions, student resources and departmental events (e.g. roundtables with invited speakers), socials, curriculum structuring, DEI efforts (e.g. antiracism readings/trainings), support for identity groups, etc. Host weekly office hours for student feedback. Sit on the departmental executive committee. Estimated 500 hours of service to the department. October 2020 – October 2021.
- Research Mentor, SMASH Rising. Co-mentoring four undergraduate students of underrepresented backgrounds in STEM. Students are exposed to and work toward identifying candidate protein biomarkers of cancer through bioinformatics approaches. Gave daily lectures on systems biology, multi-omics, mass spectrometry, data wrangling, and data visualization. Introduced Python via Google Colab notebooks. Helped form research questions culminating in a research project. June 2021 – August 2021.
- Panelist, Stanford ADVANCE - Rotations Workshop. July 9, 2021.
- Committee Member, DBDS Faculty JEDI Committee. See details below for efforts related to justice, equity, diversity, and inclusion (JEDI). September 2020 – September 2021.
- Committee Member, Student DEI Committee of the Biomedical Informatics Training Program. Student and faculty efforts to diversify admissions and culture at Stanford's Department Biomedical Data Science (DBDS). Spearheaded bi-annual departmental DEI Town Halls, antiracism book clubs, the funding of Women and Non-binary Students of Biomedical Data Science, an application review program for prospective students, and a bootcamp to onboard a diverse incoming MS and PhD student body through informatics and career development workshops. June 2020 – September 2021.
- Mentor, BIODS 360: Inclusive Mentorship in Data Science. Working with undergraduates at other institutions to introduce applied data science concepts and career paths in industry and academia. Tailoring weekly exercises to learn concepts. March 2021 – June 2021.
- Mentor, ADVANCE Undergraduate Institute (AUI). Mentored future Stanford applicants in interview preparation and application reviews. April 2021.
- Graduate Teaching Assistant, BIOMEDIN 214 (Representations & Algorithms for Computational Molecular Biology). Worked with Prof. Russ Altman to update syllabus, headed office hours and review sessions, developed exam questions and code for projects. Designed and gave two lectures: Python for bioinformatics and deep learning methods applied to biological data. Received teaching award from the Stanford Center for Professional Development (SCPD). Autumn 2019 & 2020.
- Reviewer, Deep Learning for Genomics, IEEE/ACM Transactions on Computational Biology and Bioinformatics 2020.
- Reviewer, ML4Health workshop, NeurIPS 2019.
- Panelist, AI in Radiology Research, Stanford Radiology Joint Research Retreat, 2019.
- Organizing Committee Member, Stanford Department of Biomedical Data Science (BDS) Scientific Retreat, 2019.
- Volunteer, The Early Detection of Cancer Conference, 2019
- Alumni Panelist, Mathematics Undergraduate Student Association (MUSA), UC Berkeley; May 3, 2018
- Organizing Committee Member, Biomedical Computation at Stanford (BCATS) Symposium 2018
- Research Mentor, Canary Cancer Research Education Summer Training (CREST) Program, 2018.
- Organizing Committee Member, Stanford Radiology Joint Research Retreat 2017
- Research Mentor, Canary Cancer Research Education Summer Training (CREST) Program, 2017.