Gautam Machiraju

6044 College Ave, Apt C | Oakland, CA 94618
614-599-6288 ◇ gmachi@stanford.edu ◇ gmachiraju.github.io

Education

Stanford University 2018 – Present
Ph.D. Biomedical Data Science (BDS)
Advisors: Profs. Parag Mallick (Radiology), Christopher Ré (CS; currently rotating)
Committee: Profs. Sylvia Plevritis (BDS, Radiology), Serena Yeung (BDS), James Zou (BDS)

- *Thesis topic*: development of interpretable deep learning methodologies for analyzing coarsely labeled high-resolution data, e.g. multiplexed and megapixel images and long-sequence time-series often found in biomedical settings
- *Interests:* Deep learning, multi-agent & multi-task learning, game theoretic analysis of optimization, computer vision, self-supervision and relational reasoning (e.g. encoding long-range dependencies and context), weak and/or coarse supervision, Alignment and interpretability, concept discovery

University of California, Berkeley 2012 – 2016
B.A. Applied Mathematics, Minor in Bioengineering

- *Interests*: Mathematical Biology (elected degree concentration) and numerical analysis with particular emphasis in differential equations, optimization, & algorithms

Current Projects

Thesis work:

- Gautam Machiraju, Arjun Desai, Khaled Saab, Michael Zhang, Dan Fu, Tri Dao, Christopher Ré, Parag Mallick. **Learning Long-range Dependencies and Salience from Coarse Labels with Contrastive Self-supervision**, (In preparation for *International Conference on Machine Learning*, 2023)
  This approach learns contextualized embeddings when sampling across long-range sequences that also satisfy properties of a metric space. Applications seek to efficiently learn salient objects (e.g. tumor region in images, seizure in signals) in coarsely labeled high-res data: gigapixel images (e.g. pathology) and time-series signals (e.g. ECG). Downstream models employ multi-agent training framework between classifiers operating at different scales. Interpretability assessed based on concordance between model explanations and expert annotations.

- Arjun Desai, Gautam Machiraju, Khaled Saab, Sarah Hooper, Parag Mallick, Akshay Chaudhari, Christopher Ré. **Visually Grounded Foundation Models**, (In preparation for *International Conference on Machine Learning*, 2023).
  This study extends the capabilities of today's vision-centric foundation models to learn contextual features outside of simply the foreground object. Motivated by biomedical and other human-centered vision applications such as radiology report generation and discovery of unseen biomarkers or diagnoses in medical imaging.

Non-thesis work:

- Ali Lenjani, Sarthak Consul, Turan Orenbas, Gautam Machiraju, Luke Carani, Pascal Geldsetzer. **Predicting Maternal and Child Health Indicators using Publicly Available Remote Sensing, Survey, and Geo-tagged Data**, (in preparation for *The Lancet*).
  Other applications of our vision methods focus on global health monitoring by predicting maternal and child health outcomes of Earth's remote villages through the use of Landsat satellite imagery (Google Earth Engine), government surveys (Demographic and Health Surveys Program), and other geotagging and remote sensing data.

Peer-reviewed Conference & Journal Publications

Note: preprints can be found on personal website.
Thesis work:

1. Hunter Boyce, Gautam Machiraju, Parag Mallick. **Spatial Statistics for Spatial Biology**, *Cell Systems* (In Review).

2. Gautam Machiraju, Sylvia Plevritis, Parag Mallick. **A Dataset Generation Framework for Evaluating Megapixel Image Classifiers & their Explanations**, *European Conference on Computer Vision (ECCV), 2022.*

Non-thesis work:

3. Yuanning Zheng, Jun John, Jayendra Shinde, Kevin Brennan, Hong Zheng, Gautam Machiraju, Olivier Gevaert. **An Integrative Tool for the Population-level Analysis of DNA Methylation**, (In Review).

4. Alexander Henry Thieme, Yuanning Zheng, Gautam Machiraju, et al. **Development and Evaluation of an Image-based Deep Learning Algorithm to Classify Skin Lesions from Monkeypox Virus Infection**, *Nature Medicine* (In Press, 2022).

5. 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**, *JAMIA* (2021).

6. CSBC/PS-ON Image Analysis Working Group*, Juan Carlos Vizcarra, Erik A. Burlingame, Yury Goltsev, Brian S. White, Darren Tyson, Artem Sokolov. **A Community-based Approach to Image Analysis of Cells, Tissues and Tumors**, *Computerized Medical Imaging and Graphics* (2021). *Acknowledged author within consortium

7. Gautam Machiraju, Parag Mallick, Hermann Frieboes. **Multicompartment Modeling of Protein Shedding Kinetics During Vascularized Tumor Growth**, *Nat Sci Reports* (2020).

Awards, Scholarships, & Fellowships

- Stanford Data Science Scholarship (50% support) via Stanford Data Science Institute (2021-2023)
- 2022 HAI Google Cloud Credits Award ($40K) for "Algorithmically identifying histo-molecular biomarkers of cancer progression" (2022-2023)

Invited Institute, Consortium, & National Presentations

- European Conference on Computer Vision (ECCV), 2022. Poster: **A Dataset Generation Framework for Evaluating Megapixel Image Classifiers & their Explanations**
- National Library of Medicine Informatics Trainee Conference, 2022. Focus Session presentation: **A Dataset Generation Framework for Evaluating Megapixel Image Classifiers & their Explanations**
- Canary Center at Stanford for Cancer Early Detection; April 20, 2022; Center-wide presentation: **Toward algorithmically identifying histopathological biomarkers of cancer progression**
- National Library of Medicine Informatics Trainee Conference, 2021. Focus Session presentation: **Learning multi-scale morphological features in multiplexed, megapixel microscopy images**

- Canary Center at Stanford for Cancer Early Detection; May 19, 2021; Center-wide presentation: **Learning multi-scale morphological features in multiplexed, megapixel microscopy images**
- National Library of Medicine Informatics Trainee Conference, 2020. Open Mic Session presentation: **Automated Node-Positive Classification and Feature Extraction Using Tumor Microenvironment Imaging**
- The Early Detection of Cancer Conference, 2019. Poster: **Accelerating plasma biomarker discovery with spatially-explicit mathematical modeling**
- Canary Center at Stanford for Cancer Early Detection; July 18, 2018; Center-wide presentation: **Accelerating plasma biomarker discovery with a spatially-explicit mathematical model**
- Center for Cancer Systems Biology (CCSB) Symposium, 2017. Poster: **Mathematical model of tumor cell heterogeneity and biomarker shedding origins**.
- Canary Symposium, 2017. Poster: **Mathematical model of tumor cell heterogeneity and biomarker shedding origins**
- Biomedical Computation at Stanford (BCATS) Symposium, 2017. Poster: **Mathematical model of cancer heterogeneity and biomarker shedding kinetics**

Academic Teaching, Service, & Leadership Experience

- **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 events (e.g. departmental Roundtables with invited speakers), curriculum, DEI efforts (e.g. Antiracism Book Clubs), socials, etc. Host weekly office hours for student feedback. Sit on 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. Introduce Python via Google Colab notebooks. Helped form research questions culminating in 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 *BMI 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
- **Trainee Panelist**, AI in Radiology Research, Stanford Radiology Joint Research Retreat; Oct 29, 2019
- **Student Planning Committee Member**, Stanford Department of Biomedical Data Science (BDS) Scientific Retreat 2019; Sept 27, 2019
- **Volunteer**, The Early Detection of Cancer Conference, 2019; Sept 24-25, 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. Mentoring students to aid in lab's NLP project (see below), ultimately with a submission to the Siemens Competition.
- **Organizing Committee Member**, Stanford Radiology Joint Research Retreat 2017
- **Research Mentor**, Canary Cancer Research Education Summer Training (CREST) Program, 2017. Culminated in building Natural Language Processing pipeline for PubMed data extraction, with ultimate goal of creating a public database of disease biomarkers.

Entrepreneurship Experience & Incubator Membership

- **Sutardja Center for Entrepreneurship and Technology, UC Berkeley**, January – May 2017
- **Free Ventures, UC Berkeley**, January – May 2015
- **Catalyst@Berkeley**, September – December 2014

Relevant Skills

Programming languages	Python (Pytorch, PT Lightning, Tensorflow), shell [extensive] R, MATLAB [intermediate] SQL, CUDA, C++, Java, Spark, Scala, Javascript, HTML/CSS [basic]
Other computing	cluster (Slurm), cloud (GCP, Kubernetes) workflow integrations (VSCode, W&B, Jupyter, rmate)
Natural languages	English [native], Telugu [intermediate], Spanish [basic]

Further Inquiry

For details regarding (undergraduate, post-baccalaureate, or graduate) research projects, industrial work experience, academic coursework, and entrepreneurial experience, please visit my Linkedin page (https://www.linkedin.com/in/gmachiraju) or personal website (gmachiraju.github.io).