Yasha Ektefaie

yasha_ektefaie@g.harvard.edu
yashaektefaie.com

Education

Harvard University   *Expected 11/2023*
Ph.D. Candidate in Biomedical Informatics
Advisor: Prof. Maha Farhat and Prof. Marinka Zitnik
Committee: Prof. Deborah Marks, Prof. Andrew Beam, Prof. Bryan Bryson

University of California Berkeley   *05/2020*
B.S. BioEngineering with Departmental Honors
B.S. Electrical Engineering and Computer Science with Departmental Honors
GPA: 3.811
Departmental Honors granted to top 3% of graduating undergrads, for each major

Research Experience

Harvard Medical School, *Department of Biomedical Informatics*   2020 – Present
Graduate Student Researcher
Advisors: Prof. Maha Farhat, Prof. Marinka Zitnik
Research is centered around developing machine learning methods to predict phenotypes from genotypes and understanding the ability of these methods to generalize to unseen genotypes.

Harvard Medical School, *Department of Biomedical Informatics*   2020 – 2022
Graduate Student Researcher
Advisors: Prof. Kun-Hsing Yu
Project designing CNNs to predict breast cancer phenotypes from pathology slides now published in ***NPJ Breast Cancer***.

Dascena   2020
Data Science Intern
Projects using random forests to predict stroke onset and gastrointestinal bleeding from UCSF EHR data now published in ***Frontiers in Neurology*** and in consideration in ***BMJ Open Gastroenterology***.

Verily   2019
Software Quality Intern
Project displaying metrics for diabetic retinopathy CNN deployed in low resource settings and other projects to facilitate FDA consideration of algorithm.

Harvard Medical School, *Department of Biomedical Informatics*   2018
Summer Research Intern
Advisors: Prof. Maha Farhat
Project dating phylogenetic trees using BEAST to illustrate heterogeneity of Tuberculosis epidemic now published in ***Lancet Microbe***.

University of California Berkeley, *Department of BioEngineering*   2017 – Present
Undergraduate Researcher
Advisors: Prof Adam Arkin, Dr. Lauren Lui

Project using a custom motif finding algorithm and RNA Chip-Seq data to challenge assumption of A-rich motif binding by Hfq-Crc protein complex.

Publications

Preprints & In Review

Ektefaie Y, Dasoulas G, Noori A, Farhat MR, Zitnik M. *Geometric multimodal representation learning.* arXiv:2209.03299 (2022).

Dixit A, Kagal A, **Ektefaie Y**, et al. *Modern lineages of Mycobacterium tuberculosis were recently introduced in western India and demonstrate increased transmissibility.* medRxiv 2022.01.04.22268645 (2022).

Allen A, **Ektefaie Y**, Garikipati A, Lam C, Green-Saxena A, Siefkas A, Barnes G, Handley M, Mataraso S, Hoffman J, Mao Q, Das R. *A Machine Learning Algorithm To Predict Gastrointestinal Bleeding Requiring Intervention.* Under consideration at BMJ Open Gastroenterology (2021).

Peer-Reviewed Journal Articles

Ektefaie Y, Dixit A, Freschi L, Farhat MR. *Globally diverse Mycobacterium tuberculosis resistance acquisition: a retrospective geographical and temporal analysis of whole genome sequences.* Lancet Microbe. 2021 Mar;2(3)

Ektefaie Y, Yuan W, Dillon D, Lin N, Golden J, Kohane I, Yu K. *Integrative Multiomics-Histopathology Analysis for Breast Cancer Classification.* NPJ Breast Cancer.

Maharjan J, **Ektefaie Y**, Ryan L, Mataraso S, Barnes G, Shokouhi S, Saxena A, Calvert J, Mao Q, Das R. *Enriching the Study Population for Ischemic Stroke Therapeutic Trials Using a Machine Learning Algorithm* Frontiers in Neurology. 2022 Jan.

Green, A.G., Yoon, C.H., Chen, M.L., **Ektefaie Y**, *et al.* A convolutional neural network highlights mutations relevant to antimicrobial resistance in *Mycobacterium tuberculosis*. *Nat Commun* 13, 3817 (2022).

Presentations

Invited Talks

Mutational Data Split for Machine Learning Models that Predict Phenotype from Genotype. Invited Spotlight Presentation. NHGRI Research Training & Career Development Annual Meeting. Duke University, Apr 2022

Poster Presentations

Ektefaie Y, Dixit A, Freschi L, Farhat M. *Tuberculosis resistance acquisition in space and time: an analysis of globally diverse M.tuberculosis whole genome sequences.* 50th Union World Conference on Lung Health, Oct 2019.

Dixit A, Kagal A, **Ektefaie Y**, Freschi L, Lokhande R, Groeschel M, Tornheim J, Gupte N, Pradhan N, Kadam D, Gupta A, Golub J, Mave V, Farhat M. *Modern lineages of Mycobacterium tuberculosis were recently introduced in western India and demonstrate increased transmissibility*, ID Week 2021.

Ektefaie Y, Mataraso S, Barnes G, Lynn-Palevsky A, Pellegrini E, Green-Saxena A, Hoffman J, Calvert J, Das R. *Enriching the Study Population for Ischemic Stroke Therapeutic Trials Using a Machine Learning Algorithm*, 50th Critical Care Congress.

Allen A, **Ektefaie Y**, Garikipati A, Lam C, Green-Saxena A, Siefkas A, Barnes G, Handley M, Mataraso S, Hoffman J, Mao Q, Das R. *A Machine Learning Algorithm To Predict Gastrointestinal Bleeding Requiring Intervention*, Digest Disease Week (DDW) 2021.

Ektefaie Y, Dixit A, Freschi L, Farhat M. *Tuberculosis resistance acquisition in space and time: an analysis of globally diverse M.tuberculosis whole genome sequences.* University of California Berkeley, Undergraduate Research Symposium Poster Session, Oct 2019.

Ektefaie Y, Lui L, Arkin A. *Characterization of the CrC Protein in Pseudomonas Fluorescens.* University of California Berkeley, Undergraduate Research Symposium Poster Session, Apr 2019.

Volunteering Experience

Mentoring

Alex Wei
Current Position: Master's in public health at Harvard School of Public Health   2022 – Present
Mentoring for master's thesis using graph neural networks to screen a library of chemical compounds for potential antibiotic effect on Tuberculosis.

Dasha Bykova   2021 – Present
Current Position: Masters' student at Lomonosov Moscow State University.
Awarded a grant from the Zimin Foundation SMTB Alumni Summer Research Program to do research at HMS with Maha Farhat. Mentoring her through two projects (1) visualizing and interpreting protein embeddings in Tuberculosis and the corresponding BERT model (2) using graphical neural networks to infer function of unannotated genes in Tuberculosis.

Ayush Noori   2021 – Present
Current Position: Undergraduate, Harvard University
Mentored on writing perspective about multimodal graph learning now on arxiv.

Justin Du   2021
Current Position: Undergraduate Yale University
Mentored on project using CNNs to predict Tuberculosis phenotypes from chest x-rays and CT scans.

Outreach

UC Berkeley Biomedical Engineering Society   2018 – 2019

Corporate Relations Committee Co-Chair
Lead committee of 30 undergraduates to organize industry dinner to connect members of biotech industry to underrepresented BioEngineering Students. 120 guests attended dinner, fundraised $1500 to pay for the dinner. Founded consulting program that pairs teams of student to technical industry experience. Program intended for disadvantaged students with no prior industry experience.

Startup Boston Week Conference   2021
Moderator
Panel: AI, Oh My: The Value & Limitations of Machine Learning for Your Startup
Moderated panel of 4 industry experts in ML leading question about the use of ML in startups. Panel was in front of live audience of 100. Panel was to promote machine learning to Boston startup community.

BIG GSAS Orientation Panel   2021
Panelist
Panelist to answer common questions asked about PhD at Harvard Medical School and to promote inclusive community at Harvard Medical School.

Journal Review

Reviewer: AAAI 2023, Pacific Symposium on BioComputing (PSB) 2023, ISMB 2022

Teaching

BMI 722: Topics in Translational Biomedical Informatics	2022
BMI 702: Foundations of Biomedical Informatics II	2022
BMI 707: Deep Learning for Biomedical Data	2021

For all courses taught I was a teaching fellow. I lead office hours and designed homework for class of fifty to sixty masters and medical students teaching them basics of machine learning ranging from implementing CNNs in Keras to advising novel research projects using machine learning and biomedical data (702/707) and basics of biomedical informatics from running genome-wide association studies to constructing polygenic risk scores (722). For BMI 722, I also lectured on the basics of statistical learning in biomedicine.

Honors & Awards

National Defense Science and Engineering Graduate (NDSEG) Fellow   2022
Graduate fellowship given to 2% of applicants (out of 8000 applicants) based on the merit of their proposed thesis research. Fully funds PhD for 3 years.

Bill & Melinda Gates Foundation Union World Conference Travel Award   2019
Award given to fund my travel, lodging, and expenses to attend Union World Conference in Hyderabad, India. Given top 1% of applicants to travel award based on the merit of the research.

Regents and Chancellor's Scholar at UC Berkeley   2016
"The Regents' and Chancellor's Scholarship is the most prestigious scholarship offered by UC Berkeley to entering undergraduate students." – from UC Berkeley website. Given to top 1% of

undergraduate applicants, out of 112,800 potential applications on the basis of potential to excel and serve as a leader at UC Berkeley and beyond.

EECS Honors Award   2020
Award given to EECS graduates who not only hold top GPA among all EECS undergraduates but also pursue interdisciplinary research and curriculum.

Dean's Honors List   2016, 2017
Honor awarded for a semester GPA in the top 10 percent of all College of Engineering undergraduates.

Extracurriculars

Running & Triathlon

When I am not doing research, I am an endurance runner and a triathlete. I am currently training to qualify for the Boston, New York, and Chicago Marathons and race in road races/triathlons regularly. Below are some races I recently ran:

> San Francisco Marathon, 2022
> Providence Marathon, 2022
> BAA Association 5K 2022
> BAA Association 10K 2022
> Cranberry Trifest 2022

Music

I am an avid piano player. I competed when I was younger and won various competitions listed below, but now I mostly compose original music and publish on SoundCloud (https://soundcloud.com/yashaekt).

American Fine Arts Festival: 1st Place Performance at Weil Recital Hall, Carnegie Hall, NY, 2015
Bay Area Music Association US International Competition 1st Place, 2015
United States Open Music Competition: Treasury of Romantic: 4th Place, 2015