YANIV YACOBY

Pronouns: he/him/his
yanivyacoby.github.io
yanivyacoby@g.harvard.edu
+1 (857) 636-9057

About

Research My goal is to empower individuals to effectively utilize Machine Learning by making the consequences of modeling assumptions and inference decisions transparent. Specifically, I develop deep probabilistic/Bayesian models and approximate inference methods designed for safety-critical domains, such as precision healthcare. My work has exposed failure mechanisms in a number of popular methods, and provides ways to mitigate these failures. My work has also explored how people (mis)understand popular techniques for explainable AI.

Teaching My goal is to create classroom environments and mentorship experiences that (a) build student ability to independently interrogate technological systems in societal contexts, (b) that develop effective research skills (e.g. technical reading, writing, communication), and (c) help students overcome common psychological barriers to learning (e.g. imposter phenomenon) using cohort-building and discussion of academic culture.

Education

Expected May 2023 — **Harvard University**, Cambridge, MA
Ph.D. in Computer Science. Advisor: Finale Doshi-Velez

Diploma May 2016 — **New England Conservatory of Music**, Boston, MA
‡M.M in Contemporary Improvisation. GPA: 3.96

Diploma May 2015 — **Harvard University**, Cambridge, MA
‡A.B. in Computer Science. Cum Laude in Field. GPA: 3.748 ‡Joint degree program

Research Experience

Jan 2018–Present — **Machine Learning**, Harvard University
Graduate Research Assistant, Advisor: Finale Doshi-Velez

- Representation learning and density estimation
- Uncertainty quantification
- Applications of Machine Learning in health-care: In Vitro Fertilization (IVF)
- Counterfactual generation for actionable recourse
- Evaluation of explanations systems for the Public Safety Assessment (PSA)

May 2021–Aug 2021 — **Statistical Genetics**, Microsoft Research New England
Research Intern @ Biomedical-ML Team, Advisor: Lorin Crawford

- Identified mechanism for why existing methods generalize poorly to underrepresented populations in Biobanks, and why traditional fixes do not mitigate the issue.
- Proposed a framework that easily extends existing methods to generalize well to underrepresented populations.

Sep 2017–Dec 2017 — **Accelerating Scientific Computation**, Harvard University
Graduate Research Assistant, Advisor: Margo Seltzer

- Researched methods for approximate lazy dynamic programming for accelerating scientific computation.
- Discovered latent structure in computation graphs of Gaussian Process regression useful for dynamic programming.

Jan 2015–May 2015 — **Live Audio Processing**, Harvard University
Graduate Research Assistant, Advisor: Hans Tutschku

- Designed a collection of real-time audio processing tools for marimba to enable it to participate in electronic, electroacoustic and folk music scenes.
- Tools include spectral processing, distortion effects, granular synthesis/improvisation tools, etc., written in Max/MSP (with externals in Java/JavaScript).

Sep 2013–Dec 2015 **Computer Systems**, Harvard University
Graduate Research Assistant, Advisor: Margo Seltzer

- Automatically Scalable Computation (ASC):
  - Developed algorithms and data structures for efficient lookups in a cache of gigabyte-size vectors with wildcards for ASC, an architecture that uses additional cores to automatically parallelize single-threaded programs.
  - Benchmarked methods on different distributions of cache access patterns.
- Efficient Random-Memory Access for Graph Databases
  - Studied properties of the Linux `madvise` system-call as a prefetching mechanism for SSD and NVM-based systems.
  - Evaluated its potential to be useful for random-access patterns in graph databases.

Publications

I served as a direct research mentor to the undergraduate/Master's co-authors whose names are underlined.

[1] **Y Yacoby**, J Girash, D Parkes. *Empowering First-Year Computer Science Ph.D. Students to Create a Culture that Values Community and Mental Health.* Under submission @ SIGCSE 2023.

[2] J Yao, **Y Yacoby**, B Coker, W Pan, F Doshi-Velez. *An Empirical Analysis of the Advantages of Finite vs. Infinite Width Bayesian Neural Networks. Under submission @ NeurIPS ICBINB 2022*

[3] ***Y Yacoby**, *W Pan, F Doshi-Velez. *Mitigating the Effects of Non-Identifiability on Inference for Bayesian Neural Networks with Latent Variables.* Accepted @ JMLR 2022.
- Previous version accepted @ the ICML Workshop on Uncertainty & Robustness in Deep Learning (UDL), 2019. TALK Spotlight Talk

[4] **Y Yacoby**, B Green, C Griffin, F Doshi-Velez. *"If it didn't happen, why would I change my decision?": How Judges Respond to Counterfactual Explanations for the Public Safety Assessment.* Accepted @ HCOMP 2022.
- Previous version accepted @ the CHI Workshop on Human Centered Explainable AI (HCXAI), 2022. TALK Oral Presentation

[5] **Y Yacoby**, W Pan, F Doshi-Velez. *Failures of Variational Autoencoders and Their Effects on Downstream Tasks.* Under submission @ JMLR.
- Previous version accepted @ the ICML Workshop on Uncertainty & Robustness in Deep Learning (UDL), 2020.

[6] *S Thakur, *C Lorsung, ***Y Yacoby**, F Doshi-Velez, W Pan. *Uncertainty-Aware (UNA) Bases for Deep Bayesian Regression Using Multi-Headed Auxiliary Networks.* Under submission @ JMLR.
- Previous version accepted @ the ICML Workshop on Uncertainty & Robustness in Deep Learning (UDL), 2020.

[7] T Guenais, D Vamvourellis, **Y Yacoby**, F Doshi-Velez, W Pan. *BaCOUn: Bayesian Classifiers with Out-of-Distribution Uncertainty.* ICML Workshop on Uncertainty & Robustness in Deep Learning (UDL), 2020.

[8] M Downs, J Chu, **Y Yacoby**, F Doshi-Velez, W Pan. *CRUDS: Counterfactual Recourse Using Disentangled Subspaces.* ICML Workshop on Human Interpretability in Machine Learning (WHI), 2020.

[9] **Y Yacoby**, W Pan, F Doshi-Velez. *Characterizing and Avoiding Problematic Global Optima of Variational Autoencoders.* Advances in Approximate Bayesian Inference (AABI), 2019, Proceedings of Machine Learning Research (PMLR) 118:1-17, 2020. TALK Spotlight Talk (Top 33%)

[10] D Vaughan, W Pan, **Y Yacoby**, EA Seidler, AQ Leung, F Doshi-Velez, D Sakkas. *The application of machine learning methods to evaluate predictors of live birth in programmed thaw cycles* (clinical abstract). American Society of Reproductive Medicine (ASRM), 2019.

[11] (In Preparation) Y Fouks, **Y Yacoby**, W Pan, F Doshi-Velez, D Vaughan, D Sakkas. *Pitfalls in Deploying an Electronic Medical Record-based Machine Learning Model to Predict Fertility Outcomes.*

Teaching

Sep 2021–Present **CS290A&B: Effective Research Practices & Academic Culture**, Harvard University
Creator/Developer and Co-instructor

- Created new syllabus and content to focus on the needs of entering Ph.D. students. Topics include: skill building (e.g. how to read research papers, communication in collaborative environments), soft skill building (e.g. managing advising relationships, how to support your peers), and academic culture (e.g. mental health in academia, normalizing and de-stigmatizing of mental health needs, discussion of power dynamics in scientific communities, healthy expectation setting, etc.).
- Two-semester class is mandatory for all entering computer science Ph.D. students (2021-2022: 27 students, 2022-2023: 45 students).
- Led class discussions, small group activities, panel discussions, and facilitated Q/A sessions with visiting faculty.
- Held office hours, supported students in managing advising relationships and skill building.
- Consults for the development of a similar course in the Applied Physics Ph.D. program, and for an orientation workshop series in the Institute of Applied Computational Sciences Master's program.
- Content found *here*. Paper about course found *here*.

Sep 2018–Dec 2018 **CS281: Advanced Machine Learning**, Harvard University
Teaching Fellow

- Taught sections, hosted office hours, contributed to course materials, wrote homework solutions, edited and graded exams, co-taught unit on Markov Chain Monte Carlo methods.

Sep 2015–Dec 2015 **CS61: Systems & Machine Organization**, Harvard University
Teaching Fellow

- Taught weekly section, hosted office hours, facilitated discussions in flipped class-room, validated course materials, graded exams, and designed course surveys.

Sep 2012–Dec 2012 **CS50: Intro to Computer Science I**, Harvard University
Teaching Fellow

- Taught weekly advanced section, hosted office hours, graded problem-sets and exams.

Guest Lectures:

May 2022 **CS136: Statistical Pattern Recognition**, Tufts University
Guest Lecturer on Variational Autoencoders

Oct 2021 **AC299r: Diversity, Inclusion and Leadership in Tech**, Harvard University
Guest Speaker: Discussion on academic culture and community building in graduate school

Nov 2018 **CS281: Advanced Machine Learning**, Harvard University
Guest Lecturer on Markov Chain Monte Carlo methods

Awards:

Spring 2022 **Certificate of Distinction in Teaching**, Derek Bok Center for Teaching, Harvard University
Fall 2021 **Certificate of Distinction in Teaching**, Derek Bok Center for Teaching, Harvard University

Mentoring

Sep 2022–Dec 2022 **AC297r: Computational Science & Engineering Capstone**, Harvard University
Project Mentor
- Mentored Master's students' research projects in probabilistic machine learning.

Sep 2019–May 2022 **AM207: Inference & Optimization**, Harvard University
Research Mentor
- Mentored undergraduate, Master's, and Ph.D. students enrolled in the course or continuing their projects after in research on probabilistic machine learning.

Research Mentees:

Undergrad: Jonathan Chu, Max Guo, David Ma, Zev Nicolai-Scanio, Claire Tseng, Annie Zhu
Master's: Blake Bullwinkel, Teresa Datta, Michael Downs, Théo Guénais, Molly Liu, Cooper Lorsung, Paul Tembo, Sujay Thakur, Dimitris Vamvourellis, Chenwei Wu, Ruby Zhang
Ph.D: Philipp Arens, Anna Trella

Community Building & Outreach

Jan 2021–Present **InTouch: A peer-to-peer support group graduate students**, Harvard University
Co-Leader (since Feb 2022), Peer Mentor
- Served as a peer mentor: supported students with various challenges (e.g. managing work/life balance, unhealthy advising relationships, unhealthy self-expectations, etc.).
- Organized outreach initiative to encourage students to seek support and recruit new peer mentors.
- Co-led InTouch's ongoing conversation with Harvard's SEAS about developing a guaranteed transitional funding program for Ph.D. students.
- Coordinated and hosted bi-weekly socials for community building.

Sep 2021–Dec 2021 **Ph.D. Working Group**, Harvard University
Mentor
- Helped students write personal statements, CVs, and practice interviewing in preparation for graduate school applications.
- Debunked common myths about Ph.D. programs that often provoke imposter syndrome / cause students to self-filter during the application process.

Oct 2020 **How to make the most out of your Ph.D.**, Harvard University
Workshop Creator and Organizer
- Topic: managing the multi-faceted (and often undiscussed) challenges of the Ph.D., such as managing expectations, communicating with your advisor, learning to support your peers, normalizing and removing stigma from common but difficult Ph.D. student experiences.
- Workshop consisted of presentation and small-group discussions led by senior Ph.D. students.
- Content found *here*.

Feb 2019–Feb 2020 **Women in Data Science Datathon Workshop**, Cambridge MA
Mentor
- Mentored teams, facilitated discussion, assisted in troubleshooting problems.

Industry Experience

May 2016–Aug 2016 **Uber**, San Francisco, CA
App Health Software Engineering Intern

- Built a suite of tools for accelerated prototyping of crash report classification.
- Developed an algorithm for crash report classification with 61.9% improvement over current algorithm for the android rider app codebase, and 48.9% for the iOS partner app codebase.
- Enhanced current crash report processing infrastructure to allow for manual resolution of redundant crash report classification.

May 2015–Aug 2015 **Kensho**, Cambridge, MA
Back-end Software Engineering Intern

- Created a "smart" newsfeed using natural language processing techniques to categorize CNBC articles by sector.
- Built a Go backend to allow browser tty-access into docker containers.
- Built a service for scheduling recurring scarping jobs, un-normalizing and stitching data from online trends

May 2013–Aug 2013 **Meta**, Menlo Park, CA
Infrastructure Software Engineering Intern

- Worked on tools to measure Meta's capacity for efforts in energy efficiency, disaster recovery, and site operations.
- Developed software to load-test front-end cluster machines, as well as a user interface to model cluster and machine behavior as function of load and monitor/control on-going load-tests.

May 2012–Aug 2012 **Labtiva**, Cambridge, MA
Back-end Software Engineering Intern

- Developed a data-mining algorithm for searching large sequences of multilingual text, in which content is arbitrarily interrupted by noise and white-space is arbitrarily missing. The algorithm was intended to translate highlight markings between differently formatted versions of the same PDF for a to-be released feature, estimated to be used heavily on both client and server sides.

Music

I am currently a performing musician. yanivyacoby.com

Discography:

May 2016 **The Corn Knight**
An hour-long story piece for marimba-piano duo, co-composed and performed with pianist Chase Morrin, under the guidance of Vijay Iyer. The piece is informed by a fictional narrative and composed of many short pieces, like chapters in a book.

July 2021 **All Over The Map**
A collaboration with Blue Thread, lead by vocalist Cristi Catt and flutist Nikola Radan, featuring medieval Galago-Portuguese cantigas, as well as Galician and Sephardic love songs from the Iberian Peninsula, reviving melodies from ancient manuscripts.

Jan 2020 **The Boston Imposters**
Collaboration with folk singer/songwriters The Boston Imposters (Maire Clement and Davey Harrison). Recorded on tracks "Periphery", "Mournful Dove" and "Old Sea-Walls".

Aug 2017 **Gapi**
Collaboration with pianist Chase Morrin and gayageum-player Do Yeon Kim. Recorded on track "Heung".

Current/Recent Collaborations:

Since Jan 2020 **Triga**
Trio with Eric Boodman (fiddle) and Anna Breger (baroque violin & nyckelharpa). Original and traditional music rooted in traditions from Austria to Sweden to Quebec.

Since Jan 2016 **Duo with Sunniva Brynnel (accordion & voice)**
Original and traditional Celtic music, focusing particularly on the Swedish and Irish folk music traditions.

Since Dec 2017 **Duo with Eric Boodman (fiddle)**
Original and traditional Celtic music, focusing particularly on Quebecois and Irish folk music traditions.

Since Jun 2017 **Blue Thread**
A collaboration with vocalist Cristi Catt and flutist Nikola Radan, focusing on medieval Galago-Portuguese cantigas, as well as Galician and Sephardic love songs from the Iberian Peninsula, reviving melodies from ancient manuscripts.

Since Jan 2018 **Fade Blue**
A collaboration with Julian Loida (percussion), Lily Honigberg (fiddle) and James Heazlewood-Dale (bass), focusing on the parting from the Irish and American Old-Time folk traditions.

Recent Performances (Abridged):

Mar 8 / Mar 5 2020 **Museum Concert Series of Rhode Island** / **UMass Lowell Saab Center for Portuguese Studies**
With Blue Thread "All Over the Map", showcasing global ballads migrating through centuries and cultures, with local participation each stop of the way, from Portugal to Greece, India, the Ozarks and beyond.

Jan 17-18 2020 **Boston Celtic Music Festival, Club Passim**
With fiddler Eric Boodman – original and traditional Irish and Quebecois tunes.

Oct 6 2019 **Adamant Community Club, VT**
With accordionist/vocalist Sunniva Brynnel – original and traditional Swedish and Irish tunes.

Jul 10 2019 **Fade Blue @ Watermelon Wednesdays**
With Julian Loida (percussion), Lily Honigberg (fiddle) and James Heazlewood-Dale (bass).

Feb 9 2019 **A Celtic Sojourn Live Radio Show, WGBH Studio**
With accordionist/vocalist Sunniva Brynnel – original and traditional Swedish and Irish tunes.

Jan 18 2019 **Fade Blue @ Boston Celtic Music Festival, Club Passim**
With Julian Loida (percussion), Zach Mayer (sax), Lily Honigberg and Chris Overholster (fiddles).

Sep 23 2017 **Milton.Live Easthampton Irish Festival**
With fiddler Win Horan of Solas and pianist Utsav Lal.

Aug 26-30 2017 **Blue Thread Portugal Tour**
With vocalist Cristi Catt and flutist Nikola Radan, reviving medieval Galego-Portuguese songs. Performed at the Praca do Giraldo in Evora, Centro Cultural in Cascais and the Casa-Museu Medeiros e Almeida in Lisbon.

Jul 5 2017 **Featured Showcase Artists @ Zeltsman Marimba Festival**
With pianist Chase Morrin, invited to perform original and improvised music from debut album *The Corn Knight*.

Skills & Interests

Code	Python, C, Java, Max/MSP, Go, OCaml
Web	HTML, JavaScript, CSS, SASS, Jekyll, MySQL, JQuery, Flask, ReactJS
Languages	Hebrew and English (bilingual), Chinese (beginner)
Interests	Cycling, hydroponic agriculture, cats, hiking