

# Avni Kothari

[avni510.github.io](https://avni510.github.io)  
[avni510@gmail.com](mailto:avni510@gmail.com)

EDUCATION	<b>University of California, San Diego</b> 2021 – 2023 M.S. in Computer Science Thesis: Foundations for Model-Agnostic Recourse Verification Advisors: <a href="#">Berk Ustun</a> & <a href="#">Lily Weng</a> Coursework: Machine Learning; Recommender Systems; Neural Networks & Pattern Recognition; Convex Optimization; Statistical NLP; Probabilistic Reasoning & Learning; Networking Systems
	<b>University of Texas at Austin</b> 2011 – 2016 B.A. in Mathematics & B.A. in Economics Minor in Computer Science Coursework: Databases; Programming Languages; Software Design; Real Analysis; Number Theory; Discrete Mathematics; Differential Equations; Linear Algebra & Matrix Theory; Econometrics
RESEARCH INTERESTS	Machine Learning, Algorithmic Fairness, Algorithmic Recourse, Interpretability, Natural Language Processing, Deep Learning, ML for Healthcare, Responsibly Deploying ML models in Safety Critical Settings
PAPERS	<u><a href="#">Bayesian Concept Bottleneck Models with LLM Priors</a></u> Jean Feng, Avni Kothari, Lucas Zier, Chandan Singh, Yan Shuo Tan <i>under review</i> , 2024 This work eliminates the need for human-annotated concepts by proposing a novel method to learn concepts by wrapping LLMs within a Bayesian framework. This approach is highly generalizable across various data modalities and allows for rigorous uncertainty quantification despite LLMs being prone to error and hallucinations. <u><a href="#">Prediction without Preclusion: Recourse Verification with Reachable Sets</a></u> Avni Kothari*, Bogdan Kulynych*, Lily Weng, Berk Ustun <i>ICLR – International Conference on Learning Representations, Top 5% among submissions</i> , 2024 * denotes equal contribution <u><a href="#">Bayesian Priors From Large Language Models Make Clinical Prediction Models More Interpretable</a></u> Avni Kothari, Daniel J. Bennett, Seth Goldman, Elizabeth Connelly, James D. Marks, Lucas S. Zier, Jean Feng <i>AMIA – American Medical Informatics Association, Podium Abstract</i> , 2024
POSTER PRESENTATIONS	<b>UCSF Retreat, AI Convergence: Preparing for the Age of AI</b> ; <i>San Francisco, CA</i> FEB. 2024 <b>ICML Workshop on Data-centric Machine Learning Research</b> ; <i>Honolulu, HI</i> JULY 2023 <b>ICML Workshop on Spurious Correlations, Invariance and Stability</b> ; <i>Honolulu, HI</i> JULY 2023 <b>ICML Workshop on Artificial Intelligence &amp; Human Computer Interaction</b> ; <i>Honolulu, HI</i> JULY 2023
AWARDS	DeepMind Fellow ( <a href="#">Article</a> ) 2021 – 2023 University Honors 2011 – 2014
CONTRIBUTED TALKS	<b>Bayesian Priors From LLMs Make Clinical Prediction Models More Interpretable</b> (Scheduled) NOV. 2024 AMIA Annual Symposium
WORK EXPERIENCE	<b>University of California at San Francisco</b> ; <i>San Francisco, CA</i> SEPT. 2023 – PRESENT <i>Data Scientist</i> <ul style="list-style-type: none"><li>• Researching and deploying healthcare ML models at San Francisco General Hospital under <a href="#">Jean Feng</a></li><li>• Researching and implementing methods with LLMs to align tabular machine learning models with clinical intuition for model interpretability and reliability</li><li>• Researching and evaluating using LLMs in conjunction with Bayesian methods to extract concepts from clinical notes</li><li>• Creating, evaluating, and deploying a 30-day all cause readmissions model for use at the hospital</li></ul>

- Building a data pipeline to process electronic health records from thousands of patients to make data compatible with machine learning algorithms

**Edovo; Chicago, IL**

JAN. 2020 – MAY 2021

*Software Engineer*

- Designed and developed an educational content platform to handle 700K+ requests per day
- Created a pipeline and nightly job to merge 4 billion rows of user event data in PostgreSQL
- Spearheaded team sessions to improve software development practices and adopt new frameworks

**8th Light; Chicago, IL**

AUG. 2017 – MAR. 2019

*Lead Software Engineer*

- Developed a diabetes management iOS app to connect patients with diabetic nurse specialists
- Enhanced a Java-based continuous deployment pipeline, seamlessly integrating with internal tools
- Mentored peers and residents through pair programming sessions and code reviews

*Resident Apprentice*

JAN. 2017 – AUG. 2017

- Created games and applications with a focus on Test Driven Development and SOLID Design
- Created an HTTP Server in Java without libraries for app deployment
- Gave company-wide talks on “Hashing Functions” and “Fun with Prime Numbers”

TEACHING  
EXPERIENCE

**Interpretability & Explainability in Machine Learning; UC San Diego**

SEPT. 2022 – DEC. 2022

Course taught by: [Berk Ustun](#)

Supported instruction for 30+ MS/PhD students in an introductory research course

**Differential Calculus Tutor; UT Austin**

MAY 2011 – AUG. 2013

Tutored undergraduates on limits, Riemann sum, continuity, derivatives, and differentiation rules

ACADEMIC  
SERVICE

**Reviewing; JAMA '24**

SERVICE

**Vision 1948**

MAY 2023 – PRESENT

Guest speaker in AI to educate young girls interested in STEM fields

**PenPal for the Incarcerated**

SEPT. 2020 – PRESENT

Correspond biweekly through letters and video chats with an incarcerated individual

**UCSF AI4All**

JULY 2024 – JULY 2024

Teaching assistant for machine learning assignments for high school students

**The Recyclery; Chicago, IL**

AUG. 2018 – MAY 2021

Drafted annual budget for a bike shop with 200+ customers, and assisted with repairs

SKILLS &  
INTERESTS

**Software:** Python, Java, Swift, Javascript, AWS, Elasticsearch, Elixir, SQL, Terraform

**Libraries:** Hugging Face, Pytorch, CPLEX, Numpy, Pandas, Sklearn, Redux, React

**Interests:** Cycling, Gardening, Knitting, Hiking, Swimming, Fiction