ABHINAV AGRAWAL

agrawal.abhinav1@gmail
Google Scholar
Website
+1 413-768-2256 . Bay Area, CA

ML PhD . AI Product zero-to-one . GenAI Expert . Strong communicator

I want to use my deep expertise in machine learning (research & engineering), experience in leading zero-to-one AI product development, and excellent communication skills to solve interesting human-centric AI challenges.

EDUCATION

UMass Amherst, MS and PhD in Computer Science, GPA: 3.96/4.0 — 2018 - 2024
Indian Institute of Technology Kanpur (IITK) , BTech in Electrical Engineering, GPA: 9.2/10 — 2014 - 2018

SELECTED EXPERIENCE

Research Assistant, UMass Amherst, Advisor: Justin Domke — Aug '18 - Dec '24

Accurate Bayesian inference using large-scale GPU compute and generative models — AISTATS'25 | Package

- Established the best practices for how to use **large GPU clusters and normalizing flows** to achieve faster & more accurate results than state-of-the-art MCMC methods (e.g. NUTS, HMC) on hard inference problems
- Open-sourced the developed methods in a **JAX-based Python library (vijax)**; downloaded $>$ **3.1K** times

Scalable inference in large models using amortized approximations — NeurIPS'21

- Pioneered an **amortized** approximation framework for inference in **large probabilistic models** that is **provably** as accurate as traditional approaches but scales to much larger models **(billions of variables)**
- Engineered an **input-order independent encoder**, gaining $>$ **10**$\times$ **improvement** in speed and scalability

Automated posterior sampling using generative modeling — NeurIPS'20 | Package | Coverage

- Curated the **first normalizing-flow-based variational inference (VI) black-box approach**, outperforming the then SOTA on **90% of models** in the benchmark; starting a new line of reliable, automated methods
- Open-sourced the developed methods in a Python library (vistan); downloaded more than **32K** times

AI Product Lead, SigIQ.ai (AIxEdTech start up, $10M seed) — Oct '23 - Jun '24

Exam prep AI tutor for India's toughest entrance exam—Indian Administrative Service (IAS) — Coverage

- **Led** team of designers, engineers, and AI experts to integrate cutting edge advances in **LLMs and RAGs**
- Conceptualized **user-centric LLM-based features**, growing the app from **0 to 200K+ users in** $<$ **8 months**

Research Scientist Intern, *Microsoft* — May '22 - Sep '22

- Developed an algorithm to learn **generative causal models** for computationally costly simulators
- Designed a loss criterion to learn **counterfactually accurate** models, increasing accuracy and speed

Applied Scientist Intern, *Amazon* — May '21 - Sep '21

- Implemented TensorFlow-based **learn-to-rank** methods for "Buy-it-Again" recommendations
- Incorporated the **ranking order** in the loss (ListMLE and NeuralNDCG), enhancing **metrics by** $>$ **2%**

SELECTED PUBLICATIONS

[1] **Abhinav Agrawal**, Justin Domke. Disentangling impact of capacity, objective, batchsize, estimators, and step-size on flow VI. *In,* **AISTATS**, *2025.*

[2] **Abhinav Agrawal**, Justin Domke. Amortized Variational Inference for Simple Hierarchical Distributions. *In,* **NeurIPS**, *2021.*

[3] **Abhinav Agrawal**, Daniel Sheldon, Justin Domke. Advances in Black-Box VI: normalizing flows, importance weighting, and optimization. *In* **NeurIPS**, *2020.*

SKILLS

Languages & Utilities	Python, Django, SQL, Bash, Git Docker, AWS, GCP, Postman, Supabase
Libraries	PyTorch, TensorFlow, JAX, Pyro NumPyro, transformers, vLLM, openai, llama-index, instructor, peft, pandas, streamlit

REVIEWER

NeurIPS	2019*, '20, '21, '22*, '23, '24*, '25
ICML	2020, '21, '22, '25
AISTATS	2024, '25
TMLR	2022, '23
ICLR	2021

* top reviewer