

Harvineet Singh

✉ hs3673@nyu.edu
📁 harvineet.github.io
🐙 [GitHub: harvineet](https://github.com/harvineet)

Summary

Research Scientist with a strong foundation in machine learning, causal inference, and robust AI systems. Proven expertise in developing scalable systems in industry, benchmarking AI models, and advancing state-of-the-art research. Passionate about integrating research into impactful real world applications.

Research Interests

- Areas: Machine learning, Statistics, Applications in Health
- Focus: Robustness, Causal inference, Fairness, ML governance, Benchmarking and evaluation
- Ph.D. thesis: [Fair, Robust, and Data-Efficient Machine Learning for Health](#)
Thesis committee: [Rumi Chunara](#), [Rajesh Ranganath](#), [Jennifer Hill](#), [Fei Wang](#), [Daniel B. Neill](#)

Education

- 2018–2023 **New York University**, *Center for Data Science*, USA,
Ph.D. in DATA SCIENCE GPA: 4.0/4.0.
Advisors: [Rumi Chunara](#), [Rajesh Ranganath](#)
My dissertation focuses on understanding and designing principled methods to deploy ML in the real world. Recognizing challenges along the entire ML pipeline, my work integrates contextual knowledge with causal and statistical thinking to [collect better data](#), to develop [fair algorithms](#), and to [evaluate robustness](#) to data shifts.
- 2010–2015 **Indian Institute of Technology (IIT) Delhi**, *Department of Mathematics*, India,
Integrated Master of Technology in MATHEMATICS AND COMPUTING GPA: 8.7/10.0.

Work Experience

- 2015–2018 **Adobe Research**, India.
RESEARCH ENGINEER
- Built scalable machine learning models for **customer behavior analytics** and **recommender systems**.
 - **Transferred multiple technologies** to Adobe Analytics collaborating with cross-functional teams.
 - Filed 6 patents, published 3 conference papers, presented demos at internal and external company events.

Research Experience

- 2023–Present **University of California, San Francisco**, USA.
RESEARCHER, Mentor: [Jean Feng](#)
- Developed a nonparametric framework to **diagnose performance shifts** of prediction models.
 - Evaluated clinical ML models across **multi-institutional datasets**, ensuring robustness to performance shifts.
 - Designed data collection strategies to **measure bias** in ML pipelines, aligning with ethical ML practices.
- Jun–Oct 2022 **Amazon AWS**, Germany.
RESEARCH INTERN, Mentors: [Chris Russell](#) and [Matthäus Kleindessner](#).
- Conducted a large-scale **benchmarking study** of fairness methods across tabular datasets and algorithms.
 - Explored theoretical properties of tree-based models, focusing on achieving **minmax-fairness** in practice.
- Jun–Aug 2021 **Microsoft Research**, USA.
RESEARCH INTERN, Mentors: [Emre Kiciman](#) and [Adith Swaminathan](#).
- Developed a method to optimize **contextual bandit** models for **ad ranking** pipelines.
 - Proposed a **causal feature importance** measure to enhance robustness of decision-making systems.
- Jun–Aug 2020 **Harvard University**, *Center for Research on Computation and Society*, USA.
SUMMER RESEARCH FELLOW, Mentor: [Hima Lakkaraju](#) and [Finale Doshi-Velez](#).
- Developed robust evaluation methods for **reinforcement learning** policies from offline data.
 - Integrated **adversarial learning** with causal inference to evaluate policies in non-stationary environments.
- May–Jul 2014 **Adobe Research**, India.
RESEARCH INTERN, Mentor: [Moumita Sinha](#).
- Proposed a machine learning model to predict **customer churn** from large web **clickstream data**.
 - Work **productized** in Adobe Experience Cloud and validated on online **marketing data** (US 11,170,407).

Selected Publications

- [A hierarchical decomposition for explaining ML performance discrepancies.](#)
Harvineet Singh, Fan Xia, Adarsh Subbaswamy, Alexej Gossmann, Jean Feng.
Conference on Neural Information Processing Systems (NeurIPS) 2024.
- [When do Minimax-fair Learning and Empirical Risk Minimization Coincide?](#)
Harvineet Singh, Matthäus Kleindessner, Volkan Cevher, Rumi Chunara, and Chris Russell.
International Conference on Machine Learning (ICML) 2023.
- [“Why did the Model Fail?”: Attributing Model Performance Changes to Distribution Shifts.](#)
Haoran Zhang*, Harvineet Singh*, Marzyeh Ghassemi, and Shalmali Joshi.
International Conference on Machine Learning (ICML) 2023.
- [Fairness Violations and Mitigation under Covariate Shift.](#)
Harvineet Singh, Rina Singh, Vishwali Mhasawade, and Rumi Chunara.
ACM Conference on Fairness, Accountability, and Transparency (FAccT) 2021.
- [Cascading Linear Submodular Bandits: Accounting for Position Bias and Diversity in Online Learning to Rank.](#)
Gaurush Hiranandani*, Harvineet Singh*, Prakhar Gupta*, Iftikhar A. Burhanuddin, Zheng Wen, and Branislav Kveton.
Conference on Uncertainty in Artificial Intelligence, (UAI) 2019. **Oral Presentation.**
- [Stuck? No Worries!: Task-aware Command Recommendation and Proactive Help for Analysts.](#)
Aadhavan M. Nambhi*, Bhanu P. R. Guda*, Aarsh P. Agarwal*, Gaurav Verma, Harvineet Singh and Iftikhar A. Burhanuddin.
ACM International Conference on User Modeling, Adaptation, and Personalization, (UMAP) 2019.
- [Show and Recall: Learning What Makes Videos Memorable.](#)
Sumit Shekhar, Dhruv Singal, Harvineet Singh, Manav Kedia, and Akhil Shetty.
Workshop on MBCC at IEEE International Conference on Computer Vision (ICCV) 2017.

Patents

- [Summarizing Video Content based on Memorability of the Video Content.](#) (US 10,311,913)
- [Personalized E-Learning Using A Deep-Learning-Based Knowledge Tracing And Hint-Taking Propensity Model.](#) (US 10,943,497)
- [Predicting Unsubscription of Subscribing Users.](#) (US 11,170,407)
- [End of period metric projection with intra-period alerts.](#) (US 11,205,111)
- [Intelligent customer journey mining and mapping.](#) (US 11,756,058)
- [Natural-language based Intelligent Analytics Interface.](#) (US 11,321,373)

Skills

Languages Python, R, Java, JavaScript, MATLAB.

Tools PyTorch, JAX, TensorFlow, Apache Spark, Scikit-Learn, Bash, Git.

Awards and Service

- 2023 **General Chair** for the [Machine Learning for Health Symposium](#) 2023.
- 2022 Selected for [Future Leaders Summit](#) on Responsible Data Science at University of Michigan.
- 2021 **Student Organizer** [Deep Learning Day](#) at KDD 2021.
- 2020,2024 **Mentor** [ARISE](#) at NYU, [AI4ALL](#) at UCSF to provide STEM research exposure to 9th-12th graders.
- 2018-2023 Awarded **Graduate Research Fellowship** by NYU Center for Data Science.
- 2014 Awarded **Scholarship** by Ministry of Human Resource Development for academic excellence.
- Paper Reviewer** for CS, Responsible ML, and Health venues: NeurIPS, ICLR (Highlighted Reviewer in 2022), ICML, AISTATS, FAccT, AIES, EAAMO, MLHC, CHIL, ML4H.