Ashudeep Singh

Principal Applied Scientist, Microsoft

(607)-379-7806
mail@ashudeepsingh.com
www.ashudeepsingh.com
SF Bay Area, CA

Research Interests

LLMs · Machine Learning · AI Safety & Alignment · RecSys & Information Retrieval · Responsible AI

Work Experience

Aug 2024 – Present
Principal Applied Scientist, *Microsoft AI*, Mountain View, CA.
- **Tech Lead, Agentic Multimodal Search** – Leading the development of multimodal retrieval and reasoning systems for Bing and Copilot. Designing architectures, agentic workflows and evaluations that enable low-latency retrieval and grounding across images, text, and structured data at web scale.
- Built scalable pipelines to distill foundation models into efficient, real-time variants for inference and personalization, exploring trade-offs among efficiency, capability, and safety.
- Researching practical challenges in model alignment and safety, with a focus on preserving capabilities and alignment properties during model pruning, distillation, supervised fine-tuning, and RLHF.

Aug 2021 – Aug 2024
Applied Scientist, Advanced Technologies Group, *Pinterest, Inc.*, Palo Alto, CA.
- **ML for Interactive Systems**: Applying and exploring state-of-the-art machine learning algorithms for ranking & retrieval models that learn from sequential human feedback data, spanning Graph ML, Reinforcement Learning, and sequential models such as Transformers, scaled to serve 100M+ requests per day.
- **LLM Safety & Bias Evaluation**: Pioneered company's efforts towards rigorously evaluating LLMs and image generation models for bias, safety, and fairness through red-teaming and continuous measurement and monitoring systems.
- **Responsible AI Framework Development**: Defining Pinterest's Responsible AI framework to ensure algorithmic fairness, diversity, and inclusive system design in Pinterest's recommender systems.

This work resulted in multiple successful product launches (in Pinterest Search, Related Pins, Related Products) with measurable user engagement improvements, research publications at ACM FAccT 2023 and ACM EC 2024, and a tutorial at NeurIPS 2022.

Jan – May 2020
Research Intern, *Google Brain*, New York, NY.
Safe Reinforcement Learning for Sequential Recommender Systems with **Alex Beutel**. Developed a sequential recommendation framework considering long-term user well-being and designed a novel policy gradient algorithm to optimize user engagement in the presence of safety constraints.

May – Aug 2019
Research Intern, *Microsoft Research*, Montreal, QC, Canada.
Feedback Loops and Producer-side Fairness in Recommender Systems with **Fernando Diaz** (FATE Group). Studied selection bias and exposure fairness in recommender system feedback loops.

May – Aug 2017
Research Intern, *Facebook*, Menlo Park, CA.
Active Learning for Multilabel Classification on Newsfeed with **Khalid El-Arini**. Optimized trade-off between human labeling cost and model performance, improving labeling efficiency by up to 30% for the same accuracy.

May – Aug 2016
Research Intern, *Microsoft Research Lab*, New York City, NY.
Contextual Bandits for Personalization of Notifications with **John Langford** and **Ryen White**. Personalized health app notifications using contextual bandits.

Education

2015 - 2021 **Ph.D. Computer Science**, *Cornell University*, Ithaca, NY, USA.
Title: Fairness of Exposure for Ranking Systems | Advisor: Thorsten Joachims | GPA: 4.0

2010 - 2015 **B.Tech.-M.Tech. Dual Degree**, *Indian Institute of Technology (IIT) Kanpur*, India.
Major: Computer Science and Engineering
GPA: 10.0/10.0 (M.Tech.), 9.6/10.0 (B.Tech.) | Honors: Academic Excellence Award – all years

Selected Publications

- **"Controlling Fairness and Bias in Dynamic Learning-to-Rank"** M. Morik*, A. Singh*, J. Hong, T. Joachims. *ACM SIGIR*, 2020. (*equal contribution) [**Best Paper Award**]
- **"Fairness in Ranking under Uncertainty"** A. Singh, D. Kempe, T. Joachims. *NeurIPS*, 2021.
- **"Building Healthy Recommendation Sequences for Everyone: A Safe Reinforcement Learning Approach"** A. Singh, Y. Halpern, N. Thain, K. Christakopoulou, E. H. Chi, J. Chen, A. Beutel. *FAccTRec Workshop at ACM RecSys*, 2020.
- **"Policy Learning for Fairness in Ranking"** A. Singh, T. Joachims. *NeurIPS*, 2019.
- **"Fairness of Exposure in Rankings"** A. Singh, T. Joachims. *KDD '18: The 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, 2018.
- **"Recommendations as Treatments: Debiasing Learning and Evaluation"** T. Schnabel, A. Swaminathan, A. Singh, N. Chandak, T. Joachims. *ICML*, 2016.

Complete list available at ashudeepsingh.com/publications and Google Scholar.

Awards and Recognition

2024 **Outstanding Reviewer Award**, The Web Conference.

2020 **Best Paper Award**, ACM SIGIR.

2019 **Outstanding Teaching Assistant Award**, Cornell Computer Science Department

2015 **Ranked First** in M.Tech. class (108 students), IIT Kanpur.

2011–2015 **Academic Excellence Award** (all years), IIT Kanpur.

2010–2014 **CBSE Merit Scholarship for Professional Studies** by Central Board of Secondary Education, India.

2012 **Summer Undergraduate Research Grant for Excellence (SURGE)**, granted by Dean Resource Planning and Generation, IIT Kanpur.

Technical Skills

- **Languages:** Python, C#, Java, C++, SQL
- **ML/AI:** PyTorch, TensorFlow, JAX, CUDA, Transformers, Vision Transformers, RAG, CLIP-style contrastive training, Diffusion models, LLM fine-tuning (instruction/RLHF), quantization, distillation.
- **Systems:** Docker, Kubernetes, Apache Spark, Vector databases (FAISS, Pinecone, Weaviate)
- **Cloud:** AWS, Azure, GCP, Distributed Computing

Professional Service

- Area Chair: NeurIPS (2023–2025), ICML (2022).
- Senior PC Member: ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO 2022).
- Reviewer: ICML, NeurIPS, ICLR, ACM FAccT, ACM RecSys, AAAI, TheWebConf.
- Ethics Reviewer: NeurIPS (2022–2023).
- **Tutorial** at NeurIPS 2022 on Fair and Socially Responsible ML for Recommendations, co-taught with Manish Raghavan (MIT) and Hannah Korevaar (Meta).
- **Teaching Assistant** for CS6780: Advanced Machine Learning (2019), CS4786: Machine Learning for Data Science (2016), CS4780: Machine Learning for Intelligent Systems (2015) at Cornell University.