Aniket Anand Deshmukh

CONTACT INFORMATION

Google Scholar: shorturl.at/cmxJS *Email:* aniketde15@gmail.com
Web: https://aniketde.github.io/ *Phone:* (734) 277-3529

EDUCATION

University of Michigan (UMich), Ann Arbor, MI USA CGPA: 4.0/4.0
Ph.D., Electrical and Computer Engineering *Aug'13 - Jan'19*

- Ph.D. Research Projects: Domain Generalization, Multitask Learning for Contextual Bandits, Simple Regret Minimization for Contextual Bandits.
- Graduate Student Instructor: EECS 545 Machine Learning (Fall 2016, Fall 2018)

Indian Institute of Technology Hyderabad (IIT-H), India CGPA: 8.63/10.0
Bachelor of Technology with Honors, Electrical Engineering *Jul'09 - Jul'13*

PROFESSIONAL EXPERIENCE

Tech Lead/ML Scientist, AWS AI Labs, Amazon, Santa Clara, CA, USA *Jan'23 - Present*

- Code Review using Large Language Models (LLMs): Led the development of Amazon Q's code review tool. Set technical vision and led a 6-person cross-org squad to productionize Amazon Q code-review; achieved 85% precision in offline evaluation, is now used by 10K+ developers, and flagged 1M+ issues in its first quarter. Currently working on LLM post-training techniques to finetune LLM to improve performance of code review.
- LLM pre-training for Code Generation: Advanced model-based data quality techniques for pre-training large-scale LLMs, focusing on 10-20B parameter models, to drive improvements in code generation performance and robustness.
- Personalized RAG Models: Drove S-team-level initiative: aligned leadership across 3 orgs and led 10-engineer team to ship personalized RAG for Amazon Q; boosted acceptance on 10% of queries from 45% to 65%.
- Chat for Data Preparation (AWS Low-Code/No-Code Solutions): Developed and enhanced the "Chat for Data Prep" feature in AWS Sagemaker Canvas, enabling intuitive chat-based data querying, manipulation, and visualization powered by LLMs. Designed automated query recommendations and built a testing framework to evaluate the pipeline.
- Mentor & Hiring Lead (2023–25): Ran a 150+-applicant pipeline—handled outreach, interviews, and project scoping; mentored 3 interns and 2 junior applied scientists to ship production features and co-author NeurIPS papers, converting 2 interns to full-time hires [1, 2, 3].

Senior Applied Scientist, Microsoft AI & Research, Mountain View, CA, USA *Mar'19 - Jan'23*

- Multi Media Ads (Retrieve images from general purpose corpus for Ad text): Developed an image-text multimodal model, increasing click-through rate (CTR) by 1.5% through A/B testing and outperforming OpenAI's CLIP model by 9.4% in internal evaluations. Coordinated team efforts across meetings, documentation, GPU resource allocation, and evaluation pipeline setup.
- Image attribute classification for product shopping: Designed and deployed a curriculum learning-based image classification model on Microsoft Ads shopping platform, generating $2M daily revenue. Mentored an intern to develop and deploy a clustering method, reducing image labeling costs and improving efficiency; findings were published in ECCV 2020 [4] and IJCNN 2021 [5].
- Smart Campaigns (AI powered feature to help small businesses manage their advertising): Developed a representation learning model to assess keyword-query similarity from click data, enabling over 10,000 advertisers to identify relevant keywords for their ads and improve targeting precision.

Graduate Student Research Assistant, UMich, Ann Arbor, MI, USA *Aug'13 - Jan'19*

- Contextual Bandits/Reinforcement Learning: Modeled personalized recommendation problems in a multi-task learning framework for contextual bandits [7], proving regret bounds. Developed an adaptive sensor selection framework for interplanetary spacecraft using a novel simple regret minimization technique, achieving $> 25\%$ improvement over cumulative regret-based algorithms. The project won multiple awards, including Best Paper at the ICML 2019 workshop on "Exploration in RL" [8].
- Domain Generalization (Out of Distribution Learning): Reduced the time complexity of kernel-based domain generalization algorithm from $O(n^2)$ to $O(n)$ using kernel approximation technique and proved the upper bound on the approximation error [6].

Skills

Languages	Python, SQL, Latex, HTML
Tools	PyTorch, Keras, TensorFlow, Scikit-learn, Pandas, Matlab, Azure ML, AWS, etc

Academic Service

- Outstanding Reviewer award at NeurIPS 2021, AISTATS 2022 and NeurIPS 2023 *Jun'21 - current*
- Reviewer, NeurIPS, ICML, ICLR, CVPR, AISTATS, JMLR, UAI, AAAI, etc *Jan'18 - current*
- Co-chair, KDD 2024 & 2025 Gen AI for Recommender Systems and Personalization *Jan'24 - current*
- Co-chair, TheWebConf 2023 Decision Making for IR and Recommender Systems *Oct'22 - Apr'23*
- Co-chair, ICLR 2023 Domain Generalization Workshop *Oct'22 - Apr'23*

Selected Publications

1. Subhojyoti Mukherjee, Ge Liu, Aniket Deshmukh, Anusha Lalitha, Yifei Ma, and Branislav Kveton. Experimental design for active transductive inference in large language models. *Arxiv*, 2024
2. Subhojyoti Mukherjee, Anusha Lalitha, Kousha Kalantari, Aniket Anand Deshmukh, Ge Liu, Yifei Ma, and Branislav Kveton. Optimal design for human preference elicitation. *Advances in Neural Information Processing Systems*, 37:90132–90159, 2024
3. Subhojyoti Mukherjee, Anusha Lalitha, Sailik Sengupta, Aniket Deshmukh, and Branislav Kveton. Multi-objective alignment of large language models through hypervolume maximization. *arXiv preprint arXiv:2412.05469*, 2024
4. Urun Dogan, Aniket Anand Deshmukh, Marcin Machura, and Christian Igel. Label-similarity curriculum learning. In *European Conference on Computer Vision*, pages 174–190. Springer, 2020
5. Aniket Anand Deshmukh, Jayanth Reddy Regatti, Eren Manavoglu, and Urun Dogan. Representation learning for clustering via building consensus. *Machine Learning*, 111(12):4601–4638, 2022
6. Gilles Blanchard, Aniket Anand Deshmukh, Urun Dogan, Gyemin Lee, and Clayton Scott. Domain generalization by marginal transfer learning. *Journal of Machine Learning Research*, 22(2):1–55, 2021
7. Aniket Anand Deshmukh, Urun Dogan, and Clay Scott. Multi-task learning for contextual bandits. In *Advances in Neural Information Processing Systems*, pages 4848–4856, 2017
8. Aniket Anand Deshmukh, Srinagesh Sharma, James W Cutler, Mark Moldwin, and Clayton Scott. Simple regret minimization for contextual bandits. In *Exploration in RL, ICML 2019 workshop arXiv preprint arXiv:1810.07371*, 2018
9. Aniket Anand Deshmukh, Abhimanu Kumar, Levi Boyles, Denis Charles, Eren Manavoglu, and Urun Dogan. Self-supervised contextual bandits in computer vision. *arXiv preprint arXiv:2003.08485*, 2020
10. Branislav Kveton, Boris Oreshkin, Youngsuk Park, Aniket Anand Deshmukh, and Rui Song. Online posterior sampling with a diffusion prior. *Advances in Neural Information Processing Systems*, 37:130463–130484, 2024
11. Jayanth Reddy Regatti, Aniket Anand Deshmukh, Frank Cheng, Young Hun Jung, Abhishek Gupta, and Urun Dogan. Offline rl with resource constrained online deployment. *arXiv preprint arXiv:2110.03165*, 2021