

Muhammad Ali

ali.muh@northeastern.edu

www.m-ali.org

Experience

Aug 2024 – Present Remote · Contract	Research Scientist, Human-Centered AI Apple Working on safety and human evaluations for user-facing machine learning features in Apple Services (Apple TV, Music etc.)
Nov 2023 – Jul 2024 Boston, MA	Postdoctoral Researcher Northeastern University, Institute for Experiential AI Part of the Responsible AI group. Worked on LLM evaluations for hallucinations and stereotyping; built multiple synthetic datasets for evals.
Sep 2018 – Oct 2023 Boston, MA	Graduate Research Assistant Northeastern University, Houry College of Computer Sciences Executed large-scale measurement studies of discrimination and harm in public-facing machine learning systems, with a focus on advertising platforms. Published work at CSCW, WSDM, USENIX Security and others.
May 2022 – Aug 2022 Burlington, MA	Research Intern, Machine Learning Oracle Measured linguistic biases in BERT models as a function of scale and training data type, both in the embedding-space and in downstream tasks. Project report: arXiv.
Aug 2021 – Nov 2021 Remote	Research Intern, Ethical AI Intel Worked with the AI and Analytics team to build automated model cards and explainability notebooks (with SHAP) for Intel AI's open source model repository.
Jun 2017 – Aug 2018 Saarbrücken, Germany	Research Assistant Max Planck Institute for Software Systems Executed research in algorithm auditing and using advertising data for demography. Advised by Krishna Gummadi, published work at FAccT 2018.
Aug 2015 – Mar 2016 Lahore, Pakistan	Backend Engineer Patari Music Built the foundation for the music platform's social features with a graph database; maintained and added new features to the backend API.

Education

2018 – 2023 Boston, MA	Ph.D., Northeastern University Computer Science; Advisor: Alan Mislove Outstanding Graduate Student in Research Award Selected Coursework: machine learning, parallel data processing, distributed systems
2016 – 2018 Saarbrücken, Germany	M.Sc., Saarland University Computer Science Selected Coursework: information retrieval, data mining and matrices, social network analysis, natural language processing, probabilistic graphical models
2011 – 2015 Lahore, Pakistan	B.S., National University of Computer & Emerging Sciences Computer Science

Skills

Technical	<i>Current use:</i> Python (jupyter, matplotlib, pytorch, pandas, numpy, huggingface etc.), R, JavaScript, SQL, Unix, L ^A T _E X; <i>Past use:</i> MATLAB, C, C++, Go, Scala, Spark, Hadoop, Node.js, PHP, React
Research	applied machine learning, natural language processing, large language models, explainable/responsible AI, data analysis (exploratory, descriptive), data visualization, dataset collection, algorithm auditing, writing, communication

Selected Publications [Google Scholar]

Preprint	Understanding the Interplay of Scale, Data, and Bias in Language Models: a Case Study with BERT Muhammad Ali, Swetasudha Panda, Qinlan Shen, Michael Wick, Ari Kobren arXiv, Jul 2024
USENIX Security '23	Problematic Advertising and its Disparate Exposure on Facebook Muhammad Ali, Angelica Goetzen, Alan Mislove, Elissa Redmiles, Piotr Sapiezynski USENIX Security Symposium, Anaheim, CA. Aug 2023
WSDM '21	Ad Delivery Algorithms: The Hidden Arbiters of Political Messaging Muhammad Ali*, Piotr Sapiezynski*, Aleksandra Korolova, Alan Mislove, Aaron Rieke WSDM, Online. Mar 2021
CSCW '19	Discrimination through Optimization: How Facebook's ad delivery can lead to skewed outcomes Muhammad Ali*, Piotr Sapiezynski*, Miranda Bogen, Aleksandra Korolova, Alan Mislove, Aaron Rieke CSCW, Austin, TX. Nov 2019 Diversity & Inclusion Award + Honorable Mention
FAccT '18	Potential for Discrimination in Online Targeted Advertising Till Speicher, Muhammad Ali, Giridhari Venkatadri, Filipe Ribeiro, George Arvanitakis, Fabricio Benevenuto, Krishna Gummadi, Patrick Loiseau, Alan Mislove FAccT, New York, NY. Feb 2018 Best Paper Nominee
	* equal contribution

Teaching

Summer 2023	Instructor Northeastern University CY 2550: Foundations of Cybersecurity
Fall 2022	Teaching Assistant Northeastern University CS 6140: Machine Learning, with Predrag Radjovac

Service

Program Committee	WWW 2023; FAccT 2021 — 2024; IC2S2 2023 — 2024 SoLaR @ NeurIPS 2023; ReGenAI @ CVPR 2024
External Reviewer	CSCW 2020, 2021, 2023; PETS 2022; NeurIPS 2021, 2023 (ethics reviewer)
Other	Northeastern PhD Admissions Committee, Fall 2021