Avijit Ghosh

(+1) 857-337-0180 | ghosh.a@northeastern.edu | evijit.github.io | evijit | evijit

Algorithmic Fairness | Ethical AI | Machine Learning | AI Explainability | Computational Social Science

Education

Northeastern University — Boston, MA
Ph.D. candidate in Computer Science (Advised by Dr. Christo Wilson) — *2019 - 2023 (Expected)*

Indian Institute of Technology (IIT) Kharagpur — Kharagpur, India
B.Tech. in Chemical Engineering, M.Tech in Financial Engineering, Minor in Computer Science — *2014 - 2019*

Doctoral Thesis

Algorithmic Fairness in the Real World: Challenges and Considerations — Expected defense 2023

- Social bias in machine learning algorithms is a widespread problem that has been addressed through various measures, but implementing fair machine learning systems in the real world is challenging due to issues like noisy demographic information, adversarial vulnerabilities, policy restrictions and complex interactions between humans and algorithms.
- In my thesis, I attempt to outline these problems in fair ML systems, with the aim to gain a more complete understanding of the challenges involved and to be able to provide technical and policy recommendations to overcome their real world implementation challenges.

Publications

Peer-Reviewed Conference

Queer In AI: A Case Study in Community-Led Participatory AI — FAccT '22
Anaelia Ovalle, Arjun Subramonian, Ashwin Singh, **Avijit Ghosh**, *and others* — *Chicago, Illinois*

Subverting Fair Image Search with Generative Adversarial Perturbations — FAccT '22
Avijit Ghosh, *Matthew Jagielski, Christo Wilson* — *Seoul, South Korea*

FairCanary: Rapid Continuous Explainable Fairness — AIES '22
Avijit Ghosh*, *Aalok Shanbhag*, Christo Wilson* — *Oxford, United Kingdom*

Algorithms that "Don't See Color": Comparing Biases in Lookalike and Special Ad Audiences — AIES '22
Piotr Sapiezynski, **Avijit Ghosh**, *Levi Kaplan, Alan Mislove, Aaron Rieke* — *Oxford, United Kingdom*

When Fair Ranking Meets Uncertain Inference — SIGIR '21
Avijit Ghosh, *Ritam Dutt, Christo Wilson* — *Montreal, Canada / Virtual*

Building and Auditing Fair Algorithms: A Case Study in Candidate Screening — FAccT '21
Christo Wilson, **Avijit Ghosh**, *Shan Jiang, Alan Mislove, Lewis Baker, Janelle Szary, Kelly Trindel, Frida Polli* — *Toronto, Canada / Virtual*

Public Sphere 2.0: Targeted Commenting in Online News Media — ECIR '19
Ankan Mullick, Sayan Ghosh, Ritam Dutt*,* **Avijit Ghosh***, *Abhijnan Chakrabarty* — *Cologne, Germany*

Peer-Reviewed Workshop

Can There be Art Without an Artist? — HEGM@NeurIPS '22
Avijit Ghosh, *Genoveva Fossas* — *New Orleans, USA*

Characterizing Intersectional Group Fairness with Worst-Case Comparisons — AIDBEI@AAAI '21
Avijit Ghosh, *Lea Genuit, Mary Reagan* — *Vancouver, Canada / Virtual*

Analyzing Political Advertisers' Use of Facebook's Targeting Features — Conpro@S&P '19
Avijit Ghosh, *Giridhari Venkatadri, Alan Mislove* — *San Francisco, USA*

SAVITR: A System for Real-time Location Extraction from Microblogs during Emergencies SMERP@WWW '18
Ritam Dutt, Kaustubh Hiware, **Avijit Ghosh**, *Rameshwar Bhaskaran* *Lyon, France*

WebSelect: A Research Prototype for Optimizing Ad Exposures based on Network Structure WITS'19
Avijit Ghosh, *Agam Gupta, Divya Sharma, Uttam Sarkar* *Dublin, Ireland*

Peer-Reviewed Journal

Connectedness of Markets with Heterogeneous Agents and the Information Cascades AAA'21
Avijit Ghosh, *Aditya Chourasiya, Lakshay Bansal, Abhijeet Chandra* *Journal*

Preprints and Working Manuscripts

Comparing Fair Classifiers With and Without Demographic Attributes Working Manuscript
Avijit Ghosh, *Pablo Kvitca , Christo Wilson*

Dangers of using improper datasets for fair ranking Working Manuscript
Avijit Ghosh, *Amifa Raj, Christo Wilson, Kristian Lum, Michael Ekstrand*

Unified Shapley Framework to Explain Prediction Drift Preprint
Aalok Shanbhag,* **Avijit Ghosh***, *Josh Rubin**

Supervised extraction of catchphrases from legal documents Term paper
Avijit Ghosh*, *Prerit Gupta*, Ritam Dutt, Kaustubh Hiware, Arpan Mandal, Kripabandhu Ghosh, Saptarshi Ghosh*

** Equal contribution*

Awards and Grants

2022 **Winner** Best Paper - Runner Up *Conpro'22*
2019 **Winner** Best Poster Award *ECIR'19*
2019 **Dean's Fellowship** for first Year PhD students (USD 72K) *Northeastern University*
2019 **Winner** Institute Order of Merit - Technology *IIT Kharagpur*
2018 **Winner** SGSIS Institute Challenge Grant (INR 1M) *IIT Kharagpur*
2017 **Silver Medal** Stock Market Analysis *Inter IIT Tech Meet, Kanpur*
2016 **Gold Medal** Software Development *Inter IIT Tech Meet, Mandi*
2012 **Governor's Medal** National Rank 5, ICSE Board *Government of West Bengal*
2010 **NTSE Scholar** National Talent Search Examination *NCERT*

Teaching

Algorithmic Auditing Northeastern University
Teaching Assistant for Dr. Piotr Sapiezynski *Spring 2022*

- Designing audits that measure the effects of interests and control noise sources
- Minimize potential harms of audits to all stakeholders
- Legal bounds of algorithm audits
- Beyond audits: potential harms that cannot be measured through audits

Academic Experience

Northeastern University — Boston, MA

Research Assistant at Khoury College of Computer Sciences — *Sep. 2019 – Present*

- Analyzing Fair ranking systems and showing how they fail in the presence of noisy protected attribute data. Also investigated adversarial attacks on the guaranteed fairness of retrieval systems.
- A cooperative fairness audit of the recommendation algorithm of PyMetrics, a talent matching software. Press Release.
- Investigated Facebook's Special Audiences system for opportunity advertisements and showed that the audience creation algorithm was still biased against women, seniors and minorities.
- Analyzed the ad reach and spend information obtained from Facebook's ad transparency feature and the personal targeting dataset from Propublica's Facebook ad dataset and showed that advertisers with higher budgets use more privacy sensitive targeting techniques like PII or Lookalike audiences. Findings published and presented at IEEE ConPro 2019.

LIG, University of Grenoble Alps — Grenoble, France

Visiting Researcher — *May 2019 – July 2019*

- Study of how news companies promote different items on social media, investigating possible patterns of differential information spreading using both posts and ads.
- We also discovered and reported an exposed access token bug to Facebook Bug Bounty.

IIT Kharagpur — Kharagpur, India

Undergraduate Researcher - Complex Networks Research Group — *2014 – 2019*

- Automated Extraction of Catchwords from Legal Documents using a novel NER tagger to help categorize lengthy legal texts.
- Automatically position user comments against relevant news article paragraphs. Presented at ECIR 2019.
- Savitr - A real-time location extraction system for disaster management using twitter. Presented at WWW-SMERP 2018.
- Classification and Summarization of tweets during a disaster event, presented at IBM Day 2016.

Industry Experience

Twitter — San Francisco, CA

Research Intern — *Sep – Dec 2021 and Jun – Aug 2022*

- Worked with the META (Machine Ethics, Transparency and Accountability) team at Twitter, to investigate the relationship between demography agnostic and demography dependent author impression fairness metrics at scale.
- Developed home timeline diversity metrics based on user feedback, to find balance between recommendation efficiency and fairness.

Fiddler Labs — Palo Alto, CA

Research Intern — *Oct 2020 – Apr 2021*

- Explain distributional shifts in Machine Learning model outputs by unifying Shapley based methods.
- Using optimal transport theory, proposed a threshold independent fairness metric that allows for real time explanations.
- Worked with the product team and civil rights lawyers in the deployment of Fiddler's Machine Learning model fairness dashboard. Introduced and incorporated intersectional fairness metrics in the product.

Xerox Research Centre — Bangalore, India

Research Intern — *May 2017 – July 2017*

- Implemented XTrack, a Smart Vehicle Tracking and Battery usage minimizing Algorithm, using BLE to relay GPS information.
- Proposed a method for Uber-like Surge Price Prediction using Spatio-Temporal techniques like the Neural Hawkes and Recurrent Marked Temporal Point Process. Awarded the title of Best Internship Project.

Google Summer of Code — Remote

GSoC Student at OpenMRS — *Apr 2016 - Aug 2016*

- Replaced the HTML XForms system used with native generated forms using the Forms REST Api in the android client of the Opensource Medical Record System. Added offline form saving. Configured Travis CI to automatically build and push the apk
to the play store.
- Overall, contributed 100K lines of code and became the top code contributor in the project repository.

Academic Service

2023	**The Web Conference**	*Reviewer*
2022	**ACM Conference on Fairness, Accountability, and Transparency**	*Reviewer*
2022	**AAAI/ACM Conference on AI, Ethics, and Society**	*Reviewer*
2022	**Conference on Neural Information Processing Systems**	*Reviewer*
2022	**Conference on Empirical Methods in Natural Language Processing**	*Reviewer*
2021	**Conference on Neural Information Processing Systems**	*Reviewer*

Outreach and Leadership

2022	**FAccT CRAFT Workshop on Identifying Queer Harms as a bias bounty with Queer in AI**	*Organizer*
2021	**SIGIR Queer in AI social with speaker panel on queer stereotypes in web search**	*Organizer*
2017	**Kharagpur Winter Of Code**	*Founder*
2016	**Kharagpur Open Source Society**	*Founder*

Speaking Engagements

2023	**Can There be AI Art Without an Artist?**	*South By Southwest (SXSW)*
2022	**Proxies for bias monitoring: Ethics workshop**	*Centre for Data Ethics and Innovation, UK*
2022	**Subverting Fair Image Search with Generative Adversarial Perturbations** as part of the 'Celebrating Young Researchers' event	*Trustworthy ML Initiative*

Press Mentions

2021	**NYC aims to be first to rein in AI hiring tools**	*Associated Press*
2021	**Auditors are testing hiring algorithms for bias, but there's no easy fix**	*MIT Technology Review*
2021	**New York City Proposes Regulating Algorithms Used in Hiring**	*Wired*
2021	**Supporting Responsible Use of AI and Equitable Outcomes in Financial Services**	*The Federal Reserve*
2019	**Facebook Ads Can Still Discriminate Against Women and Older Workers, Despite a Civil Rights Settlement**	*Propublica*
2019	**Facebook Agreed Not to Let Its Ads Discriminate. But They Still Can.**	*Mother Jones*

Technical Skills

Languages: Python, Java, C, R, Bash, SQL, HTML/CSS, JavaScript, Matlab

Tools & Frameworks: Git, TravisCI, Pytorch, Keras, TensorFlow, Docker, AWS, Sagemaker, Google Cloud ML, Android

Machine Learning: Transformers/BERT, ResNets, Adversarial examples, Fairness, Explainability