

Swapneel Mehta

mehtaver.se | swapneel@onefact.org | LinkedIn

EDUCATION

NEW YORK UNIVERSITY

PHD IN DATA SCIENCE

(exp) Aug 2023 | New York, NY

Cum. GPA: 3.8/4

MUMBAI UNIVERSITY

BENG IN COMPUTER SCIENCE

June 2018 | Mumbai, India

Cum. GPA: 9.23/10

COURSEWORK

DOCTORAL

Machine Learning

Causal Inference

Inference and Representation

Math Tools for Data Science

Big Data

Probability and Statistics

(Instructor, funded by Google Research)

Unicode ML Summer School

UNDERGRADUATE

Artificial Intelligence

Graph Theory

Machine Learning

Operating Systems

Database Management Systems

Computer Networks + Labs

Computer Graphics + Labs

(Co-founder, FOSS Training Program)

Unicode

AWARDS

- 2023 **Google Research Innovator**
- 2023 **WSDM Smart City Challenge**
- 2020 **IRIS-HEP Fellowship**
- 2018 **Oxford MLHEP Challenge**
- 2018 **India's Top 33 Engineers, Economic Times**

SKILLS

Python, Pytorch • MySQL • Bash • R
GCP • BigQuery • Unix/Linux • Git
S3 • Docker • LAMP • SLURM
HTML/CSS • Javascript • C++ • Java

LINKS

Github:// [swapneelm](#)

Twitter:// [@swapneel_mehta](#)

YouTube:// [swapneel_mehta](#)

EXPERIENCE

FOUNDER | SIMPPL

Dec 2021 - Present | New York, NY

- Developing tools to identify multimodal disinformation operations by coordinated networks on social media platforms. Global south misinformation.
- Civic Integrity Products: **Parrot Report**, SimPPL
- Clients: VTDigger, Yale Daily News, Sunday Times, Deutsche Welle ([summary](#))
- Selected as a **Google Research Innovator**; awarded USD 35,000 in grants from AWS, Google, the **Wikimedia Foundation**, and the **NYC Media Lab**

CERN, CMS EXPERIMENT | TSP RESEARCHER (ML)

March 2018 - July 2019 | Geneva, Switzerland

- Developed graph neural networks for particle track reconstruction at the LHC
- Productionized the **DeepJet** ML framework; won awards from Google, Oxford.

NYU, CSMAP | GRADUATE RESEARCHER

Feb 2021 - Present | New York, NY

- Identify the causal effect of feedback on low quality news posts on Reddit.
- Identify the causal impact of Twitter's warning labels on Trump's tweets (**Stanford Trust & Safety Research Conference**).
- Infectious disease modeling with probabilistic programs (**MIT PROBPLOG '21**).

INTERNSHIPS

TWITTER, CIVIC INTEGRITY | MACHINE LEARNING ENGINEER

June 2022 - Aug 2022 | New York, NY

- Improved the precision of their civic and health misinfo. classifier (BigQuery ML) by **20%** scoring 1M+ tweets / week. NLP; recommender systems.

OXFORD, TORR VISION GROUP | VISITING RESEARCHER

Aug 2021 - June 2022 | Oxford, UK

- Generative modeling of adversarial agent behaviors on Reddit recommendations presented in workshops at **ICML** and **DEFCON**, 2022.

ADOBE RESEARCH | DATA SCIENCE RESEARCH INTERN

May 2020 - Dec 2020 | New York, NY

- Trending Hashtag Recommendation for Videos; NLP, language models, data mining, weakly-supervised learning, multimodal graph neural networks.
- First-author paper accepted at IEEE ISM; U.S. Patent App P10634-US.

IIT BOMBAY, MICROSOFT | RESEARCH INTERN

May 2016 - May 2017 | New York, NY

- **Prof. S. Sahasrabudhe**, Automated evaluation of Blender 3D Modeling assignments (ICALT 2018); edX Prize Finalist, 2019.
- **Microsoft Research**: Google Analytics for the **Lokacart** D2C app for farmers.

SMOKESCREEN (ACQ. ZSCALER) | CYBERSECURITY / ML INTERN

Nov 2016 - March 2017 | New York, NY

- Reducing the footprint of network scans and building Markov-chain based hostname generation for intelligent decoys.
- Automated deployment of NASSCOM-DSCI awarded IllusionBlack software.

REVIEWER

2023 AABI
2023 ICLR, AI4ABM Workshop
2023 AI+Local News Challenge
2022 AISTATS
2022 ICML, AI4ABM Workshop
2022 NeurIPS, WBRC
2019 Elsevier JPDC

ORGANIZING

2023 AI, Misinfo. & Policy Seminar
2023 ICLR AI4ABM Workshop
2023 Misinfo. Village, DEFCON
2023 NYU AI School 2023
2022 NeurIPS WBRC
2022 NYU AI School 2022
2022 AI for Everyone, C+J 2022
2021 NYU AI School 2021
2021 Unicode ML Summer School

TALKS

2023 McKinsey Guest Talk
2023 MIDAS Future Leaders Summit (Responsible AI)
2023 CCNY SASE "TED Talks"
2023 MozFest by Mozilla
2023 NYU Data Science Seminar
2023 Deutsche Welle (multiple)
2023 NYC Media Lab
2022 Graphika
2022 Truth and Trust Online Conf.
2022 Twitter, Health ML
2022 MIT Media Lab
2022 Stanford, T&S Research Conf.
2022 NYU Graduate Seminar
2022 Misinfo. Village, DEFCON
2022 Oxford, Torr Vision Group
2021 Twitter, Cortex
2021 Facebook, Probability
2021 Flatiron Comp. Bio. Group
2020 NYC AI Workshop
2020 NYU Data Science
2018 Mumbai University (multiple)

PUBLICATIONS

- [1] S. Mehta, J. Bisbee, Z. Sanderson, R. Bonneau, J. Tucker, and J. Nagler. **Identifying the Causal Effects of Twitter's Interventions on Trump's Tweets**. *Working Paper*, 2023.
- [2] S. Mehta, K. Aslett, R. Bonneau, J. Tucker, and J. Nagler. **Identifying the Causal Effects of Feedback Reduction on Posts containing Low-Quality News on Reddit**. *Working Paper*, 2023.
- [3] S. Mehta, A. G. Baydin, R. Bonneau, J. Nagler, and P. Torr. **Estimating the Impact of Coordinated Inauthentic Behavior on Content Recommendations in Social Networks**. In *ICML 2022 Workshop AI for Agent-Based Modelling*, 2022.
- [4] S. Mehta, S. Sarkhel, X. Chen, S. Mitra, V. Swaminathan, R. Rossi, A. Aminian, H. Guo, and K. Garg. **Open-domain Trending Hashtag Recommendation for Videos**. In *IEEE International Symposium on Multimedia (ISM)*, 2021.
- [5] S. Mehta and N. Kasmanoff. **Covid-19 Modeling and Control via Policy Interventions**. In *International Conference on Probabilistic Programming (PROBPROG)*, pages https-probprog, 2021.
- [6] D. Gandhi, R. Jain, J. Gala, J. Lalwani, and S. Mehta. Expanding access to ml research through student-led collaboratives. In *NeurIPS 2022 Workshop on Broadening Research Collaborations*, 2022.
- [7] A. Ganeriwala and S. Mehta. Application of ols regression and var models to analyse the economies of varying political regimes. *Journal of Student Research*, 10(4), 2021.
- [8] S. Mehta, P. Kothuri, and D. L. Garcia. A big data architecture for log data storage and analysis. In *Integrated Intelligent Computing, Communication and Security*, pages 201–209. Springer, Singapore, 2018.
- [9] S. Mehta, C. Raman, N. Ayer, and S. Sahasrabudhe. Auto-grading for 3d modeling assignments in moocs. In *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*, pages 51–53. IEEE, 2018.
- [10] S. Mehta, P. Kothuri, and D. L. Garcia. Anomaly detection for network connection logs. *arXiv preprint arXiv:1812.01941*, 2018.
- [11] S. Mehta, M. Verzetti, J. Kieseler, and M. Stoye. Deepjet: A machine learning environment for high-energy physics. 2018.
- [12] M. Feickert, M. Paganini, S. Mehta, K. Cranmer, J. Low, D. Guest, B. Stienen, S. Templ, R. Di Sipio, M. Stoye, et al. iml-wg/hep-ml-resources: 2019. 2020.
- [13] J. Kieseler, M. Stoye, M. Verzetti, P. Silva, S. Mehta, A. Stakia, Y. Iiyama, E. Bols, S. Qasim, H. Kirschenmann, et al. Deepjetcore. *Zenodo, doi*, 10, 2020.

GRANTS

2023	Wikimedia Foundation 'WikiCred' Grant	USD 10,000
2023	Deutsche Welle	EUR 4,950
2022	GCP, AWS, Algovera Research Grants	USD 9,000
2022	JournalismAI	USD 2,000
2022	The Sunday Times	USD 2,000
2021	UK Research & Innovation Grant (w/ Oxford TVG)	GBP 23,000
2021	NYC Media Lab AI + Local News Challenge	USD 7,500
2021	Google Research India	INR 85,000