Sanghamitra Dutta

Website: sites.google.com/site/sanghamitraweb Google Scholar: bit.ly/SanghamitraGoogleScholar

EMPLOYMENT University of Maryland, College Park Assistant Professor in Electrical and Computer Engineering Affiliate Appointments: Center for Machine Learning (CML) University of Maryland Institute for Advanced Computer Studies (UMIACS) **Department of Computer Science** Values-Centered Artificial Intelligence (VCAI) JP Morgan Chase AI Research, New York July 2021-July 2022

Research Scientist in Explainable AI Center of Excellence (XAI CoE)

EDUCATION

Carnegie Mellon University

Ph.D. in Electrical and Computer Engineering Advisor: Prof. Pulkit Grover Thesis: Strategies for Fair, Explainable, and Reliable Machine Learning Using Information Theory Committee: Pulkit Grover, Kush Varshney, Jose Moura, Anupam Datta, Alexandra Chouldechova

Indian Institute of Technology, Kharagpur B. Tech. in Electronics and Electrical Communication Engineering Project Advisor: Prof. Arijit De

RESEARCH INTERESTS

The goal of my research is to build the foundations of reliable and trustworthy machine learning and carry them all the way to practice so that AI can truly bring about social good. I am particularly interested in addressing the challenges around fairness, explainability, and privacy, by bringing in a novel foundational perspective using tools from information theory, probability, optimization, and causality. I am quite excited by the fundamental and transdisciplinary challenges that arise in my research, be it ethical, legal, or computational. I am driven towards addressing these challenges by collaborating with experts across computer systems, finance, policy and law. My work has appeared in machine learning conferences, namely, NeurIPS, ICML, ICLR, AAAI, AISTATS, FAccT, AIES, as well as, information-theory venues, namely, ISIT and IEEE Transactions on Information Theory, featured in New Scientist and Montreal AI Ethics Brief and also been adopted as part of the fair lending model review at JPMorgan.

AWARDS, FELLOWSHIPS, HONORS (in reverse chronological order)

Post-Graduation	
George Corcoran Memorial Award.	2024
NSF CAREER Award.	2024
JPMorgan Faculty Award.	2023
Northrop Grumman Seed Grant.	2022

A. G. Milnes Thesis Award

Aug 2015 - May 2021

July 2011 - Aug 2015 **Best Project Award**

Aug 2022 - present

Simon's Institute Fellowship for Program in Causality. A.G. Milnes Outstanding Thesis Award at CMU.	2021 2021
Graduate CMU Cylab 2020 Presidential Fellowship. Oral Paper Presentation at AAAI. K&L Gates 2019 Presidential Fellowship in Ethics and Computational Technologies Axel Berny 2019 Presidential Graduate Fellowship. Selected to Participate in Women in Data Science and Mathematics Workshop (Wi Tan Endowed Graduate Fellowship. Prabhu and Poonam Goel Graduate Fellowship. Carnegie Institute of Technology Dean's Fellowship.	2019
Undergraduate Nilanjan Ganguly Memorial Award for Best Undergraduate Project. HONDA Young Engineer and Scientist Award. MITACS Globalink Research Award. DAAD Wise Scholarship. Secured 99.87 percentile in IITJEE, State Rank 1 in AIEEE, Rank 6 in WBJEE. Qualified Entrance Exam for Indian Statistical Institute (ISI). KVPY Fellowship from the Government of India. Regional Mathematical Olympiad (RMO).	2015 2014 2014 2014 2011 2011 2011 2010
INTERNSHIPS	
Dataminr Manager(s): Alejandro Jaimes, Joel Tetreault Mentors: Liang Ma, Tanay Saha Project: Event Extraction for Natural Language Processing Using Graph Neural Net	June 2020 - Aug 2020 vorks
IBM TJ Watson Research Center Manager: Kush Varshney Mentors: Dennis Wei, Hazar Yueksel, Pin-Yu Chen, Sijia Liu Project: An Information-Theoretic Perspective on the Relationship Between Accurac	May 2019 - Aug 2019 y and Fairness
IBM TJ Watson Research Center Manager: Priya Nagpurkar Mentors: Prof. Gauri Joshi, Parijat Dube Project: Quantifying Error-Runtime Trade-Offs for Distributed Stochastic Gradient P	May 2017 - Aug 2017 Descent
University of Alberta, Edmonton Mentors: Prof. Mrinal Mandal, Prof. Karumudi Rambabu Project: Improving Buried Object Imaging Using Ultra Wide Band (UWB) Radar	May 2014 - Jul 2014

PUBLICATIONS

Selected Conference Papers

- [C20] P. Dissanayake and S. Dutta. "Model Reconstruction Using Counterfactual Explanations: A Perspective From Polytope Theory". In: Advances in Neural Information Processing Systems (NeurIPS). 2024.
- [C19] F. Hamman and **S. Dutta**. "Demystifying Local and Global Fairness Trade-offs in Federated Learning Using Partial Information Decomposition". In: *International Conference on Learning Representations (ICLR)*. 2024.
- [C18] F. Hamman and **S. Dutta**. "A Unified View of Group Fairness Tradeoffs Using Partial Information Decomposition". In: *IEEE International Symposium on Information Theory (ISIT)*. 2024.
- [C17] F. Hamman, E. Noorani, S. Mishra, D. Magazzeni, and S. Dutta. "Robust Counterfactual Explanations for Neural Networks with Probabilistic Guarantees". In: *International Conference on Machine Learning* (ICML). July 2023.
- [C16] F. Hamman, J. Chen, and S. Dutta. "Can Querying for Bias Leak Protected Attributes? Achieving Privacy With Smooth Sensitivity". In: ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT). June 2023.
- [C15] S. Sharma, S. Dutta, E. Albini, F. Lecue, D. Magazzeni, and M. Veloso. "REFRESH: Responsible and Efficient Feature Reselection guided by SHAP values". In: AAAI/ACM Conference on AI, Ethics, and Society (AIES). Aug. 2023.
- [C14] S. Garg, S. Dutta, M. Dalirrooyfard, A. Schneider, and Y. Nevmyvaka. "In- or Out-of-Distribution Detection via Dual Divergence Estimation". In: *Conference on Uncertainty in Artificial Intelligence* (UAI). Aug. 2023.
- [C13] P. Mathur, A. T. Neerkaje, M. Chhibber, R. Sawhney, F. Guo, F. Dernoncourt, S. Dutta, and D. Manocha. "MONOPOLY: Financial Prediction from MONetary POLicY Conference Videos Using Multimodal Cues". In: ACM Multimedia (ACM-MM). Oct. 2022.
- [C12] **S. Dutta**, J. Long, S. Mishra, C. Tilli, and D. Magazzeni. "Robust Counterfactual Explanations for Tree-Based Ensembles". In: *International Conference on Machine Learning (ICML)*. July 2022.
- [C11] P. Venkatesh, S. Dutta, N. Mehta, and P. Grover. "Can Information Flows Suggest Targets for Interventions in Neural Circuits?" In: *Advances in Neural Information Processing Systems (NeurIPS)*. Vol. 34. Dec. 2021, pp. 3149–3162.
- [C10] S. Dutta, D. Wei, H. Yueksel, P. Y. Chen, S. Liu, and K. Varshney. "Is There a Trade-Off Between Fairness and Accuracy? A Perspective Using Mismatched Hypothesis Testing". In: *International Conference on Machine Learning* (*ICML*). *Featured in New Scientist*. July 2020, pp. 2803–2813.
- [C9] P. Venkatesh, **S. Dutta**, and P. Grover. "How else can we define Information Flow in Neural Circuits?" In: *IEEE International Symposium on Information Theory* (**ISIT**). June 2020, pp. 2879–2884.
- [C8] S. Dutta, P. Venkatesh, P. Mardziel, A. Datta, and P. Grover. "An Information-Theoretic Quantification of Discrimination with Exempt Features". In: AAAI Conference on Artificial Intelligence (AAAI, Oral Presentation). Feb. 2020, pp. 3825–3833.
- [C7] P. Venkatesh, S. Dutta, and P. Grover. "How should we define Information Flow in Neural Circuits".In: *IEEE International Symposium on Information Theory (ISIT)*. July 2019, pp. 176–180.
- [C6] U. Sheth, S. Dutta, M. Chaudhari, H. Jeong, Y. Yang, J. Kohonen, T. Roos, and P. Grover. "An Application of Storage-Optimal MatDot Codes for Coded Matrix Multiplication: Fast K-Nearest Neighbors Estimation". In: *IEEE International Conference on Big Data* (*IEEE BigData*). Dec. 2018, pp. 1113–1120.

- [C5] S. Dutta*, Z. Bai*, H. Jeong, T. M. Low, and P. Grover. "A Unified Coded Deep Neural Network Training Strategy based on Generalized PolyDot codes". In: *IEEE International Symposium on Information Theory (ISIT)*. June 2018, pp. 1585–1589.
- [C4] S. Dutta, G. Joshi, P. Dube, S. Ghosh, and P. Nagpurkar. "Slow and Stale Gradients Can Win the Race: Error-Runtime Trade-offs in Distributed SGD". In: *International Conference on Artificial Intelligence and Statistics (AISTATS)*. Apr. 2018, pp. 803–812.
- [C3] S. Dutta, V. Cadambe, and P. Grover. "Coded convolution for parallel and distributed computing within a deadline". In: *IEEE International Symposium on Information Theory (ISIT)*. June 2017, pp. 2403–2407.
- [C2] S. Dutta, V. Cadambe, and P. Grover. "Short-Dot": Computing Large Linear Transforms Distributedly Using Coded Short Dot Products". In: *Advances on Neural Information Processing Systems* (*NeurIPS*). Dec. 2016, pp. 2100–2108.
- [C1] S. Dutta and P. Grover. "Adaptivity provably helps: information-theoretic limits on l_0 cost of nonadaptive sensing". In: *IEEE International Symposium on Information Theory (ISIT)*. July 2016, pp. 1431–1435.

Selected Journal Papers

- [J10] F. Hamman, E. Noorani, S. Mishra, D. Magazzeni, and S. Dutta. "Robust Algorithmic Recourse Under Model Multiplicity With Probabilistic Guarantees". In: *IEEE Journal on Selected Areas in Information Theory (JSAIT)* (2024).
- [J9] A. K. Veldanda, I. Brugere, S. Dutta, A. Mishler, and S. Garg. "Hyper-parameter Tuning for Fair Classification without Sensitive Attribute Access". In: *Transactions on Machine Learning Research* (*TMLR*) (2024).
- [J8] **S. Dutta** and F. Hamman. "A Review of Partial Information Decomposition in Algorithmic Fairness and Explainability". In: *Entropy* 25.5 (2023).
- [J7] A. K. Veldanda, I. Brugere, J. Chen, S. Dutta, A. Mishler, and S. Garg. "Fairness via In-Processing in the Over-parameterized Regime: A Cautionary Tale". In: *Transactions on Machine Learning Research (TMLR)* (2023).
- [J6] **S. Dutta**, P. Venkatesh, P. Mardziel, A. Datta, and P. Grover. "Fairness under Feature Exemptions: Counterfactual and Observational Measures". In: *IEEE Transactions on Information Theory* (2021).
- [J5] **S. Dutta**, J. Wang, and G. Joshi. "Slow and stale gradients can win the race". In: *IEEE Journal on Selected Areas in Information Theory* (2021).
- [J4] S. Dutta*, H. Jeong*, Y. Yang*, V. Cadambe, T. M. Low, and P. Grover. "Addressing Unreliability in Emerging Devices and Non-von Neumann Architectures Using Coded Computing". In: *Proceedings of the IEEE* 108.8 (Aug. 2020).
- [J3] P. Venkatesh, **S. Dutta**, and P. Grover. "Information flow in computational systems". In: *IEEE Transactions on Information Theory* 66.9 (2020).
- [J2] S. Dutta*, M. Fahim*, F. Haddadpour*, H. Jeong*, V. Cadambe, and P. Grover. "On the Optimal Recovery Threshold of Coded Matrix Multiplication". In: *IEEE Transactions on Information Theory* 66.1 (2020).
- [J1] **S. Dutta**, V. Cadambe, and P. Grover. ""Short-Dot": Computing Large Linear Transforms Distributedly Using Coded Short Dot Products". In: *IEEE Transactions on Information Theory* 65.10 (2019).

Other Papers (Peer-Reviewed Workshops/Conferences/Journals)

- [W11] B. Halder, F. Hamman, P. Dissanayake, Q. Zhang, I. Sucholutsky, and **S. Dutta**. "Quantifying Spuriousness of Biased Datasets Using Partial Information Decomposition". In: *ICML Workshop on Data-centric Machine Learning Research (DMLR): Datasets for Foundation Models* (2024).
- [W10] F. Hamman and **S. Dutta**. "Demystifying Local and Global Fairness Trade-offs in Federated Learning Using Partial Information Decomposition". In: *ICML Workshop on Federated Learning and Analytics in Practice* (2023).
- [W9] F. Hamman, J. Chen, and **S. Dutta**. "Can Querying for Bias Leak Protected Attributes? Achieving Privacy With Smooth Sensitivity". In: *NeurIPS Workshop on Algorithmic Fairness through the Lens of Causality and Privacy* (2022).
- [W8] S. Mishra, **S. Dutta**, J. Long, and D. Magazzeni. "A Survey on the Robustness of Feature Importance and Counterfactual Explanations". In: *Explainability in Finance Workshop at ICAIF* (2021).
- [W7] **S. Dutta**, P. Venkatesh, and P. Grover. "Quantifying Feature Contributions to Overall Disparity Using Information Theory". In: *AAAI Workshop on Information-Theoretic Methods for Causal Inference and Discovery* (2021).
- [W6] C. Jiang^{*}, B. Wu^{*}, **S. Dutta**, and P. Grover. "An Information-Theoretic Measure for Enabling Category Exemptions with an Application to Filter Bubbles". In: *BIAS Workshop at ECIR* (2021).
- [W5] **S. Dutta**, L. Ma, T. Saha, A. Jaimes, and J. Tetreault. "GTN-ED: Event Detection Using Graph Transformer Networks". In: *TextGraphs Workshop at NAACL* (2021).
- [W4] S. Dutta, Z. Bai, T. M. Low, and P. Grover. "CodeNet: Training Large Scale Neural Networks in Presence of Soft-Errors". In: Coding Theory For Large-scale Machine Learning Workshop at ICML (CodML Workshop, ICML, Spotlight Presentation). June 2019.
- [W3] M. Fahim, H. Jeong, F. Haddadpour, **S. Dutta**, V. Cadambe, and P Grover. "On the optimal recovery threshold of coded matrix multiplication". In: *Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. Oct. 2017.
- [W2] S. Dutta*, Y. Yang*, N. Wang, E. Pop, V. Cadambe, and P. Grover. "Reliable Matrix Multiplication using Error-prone Dot-product Nanofunctions with an application to logistic regression". In: Semiconductor Research Corporation TECHCON. Sept. 2016.
- [W1] **S. Dutta** and A. De. "Sparse UltraWideBand Radar Imaging in a Locally Adapting Matching Pursuit (LAMP) Framework". In: *IEEE International Radar Conference*. May 2015.

[* denotes equal contribution]

News/Discussion on My Research

· Research on fairness tradeoffs in federated learning featured in Montreal AI Ethics Brief (see article).

 \cdot Research on algorithmic fairness adopted as part of the fair lending model review at JP Morgan and is being patented.

· Participant in the Causality Program at the Simon's Institute in Spring 2022.

• Research on accuracy-fairness trade-offs and creation of ideal datasets featured in an article titled "Teaching an AI to be less biased doesn't have to make it less accurate" in New Scientist.

• Research on fairness under feature exemptions featured in an article titled "Managing necessary bias in AI" in CMU News on Engineering Research.

· Research on coded computing for deep neural network training featured in an article titled, "Strength

training deep neural networks," in CMU News on Engineering Research.

· Research on adaptive scheduling for distributed SGD patented at IBM Research.

 \cdot Research on coded computing (including Short-Dot codes and MatDot codes) addresses problems in algorithm-based fault-tolerance that have been open for several decades and thus have received substantial attention from across communities.

Patents/Invention Disclosures

- [P4] S. Mishra, F. Lecue, C. Tilli, D. Magazzeni, **S. Dutta**, and J. Long. *Method And System For Computing Unstability Factors In Predictive Model*. US App. 18/112,272. 2023.
- [P3] Ivan Brugere et al. *Method And System For Improving Model Fairness By Using Explainability Techniques.* Invention Disclosure. 2022.
- [P2] P. Grover, H. Jeong, Y. Yang, S. Dutta, Z. Bai, T. M. Low, M. Fahim, F. Haddadpour, and V. Cadambe. Coded computation strategies for distributed matrix-matrix and matrix-vector products. US App. 16/588,990. Apr. 2020.
- [P1] P. Dube, **S. Dutta**, G. Joshi, and P. Nagpurkar. *Adaptive learning rate schedule in distributed stochastic gradient descent*. US Patent 15/938,830. Oct. 2019.

INVITED TALKS

· CMU Cylab.	2024
· ISIT Workshop on Information-Theoretic Methods for Trustworthy Machine Learning.	2024
· Values-Centered AI Seminar (VCAI, UMD).	2024
• Information Theory and Applications Workshop (ITA).	2024
\cdot Math Machine Learning Seminar (MPI + UCLA).	2023
· Simons Institute Workshop.	2023
· CISS Conference, John Hopkins.	2023
· Penn State.	2023
· Simons Institute.	2023
· Laboratory for Telecommunication Science (UMD).	2022
· SPS Seminar, IEEE Washington Chapter.	2022
· CCSP Seminar, UMD.	2022
· INFORMS Annual Meeting.	2022
· Bocconi University.	2022
· Simons Institute.	2022
· University of Maryland.	2021
· University of Michigan.	2021
· University of Southern California.	2021
· Microsoft Rising Stars.	2021
· JP Morgan Chase AI Research.	2021
· Scripps College.	2021
· New York University.	2021
· Harvard University.	2021
· Fair February, Brown University.	2021
· Data Science Seminar, University of Utah.	2021
· Microsoft Research.	2021
· Google Research.	2020

· Experian DataLabs.	2020
· CMU AI for Social Good Symposium.	2020
. Women in Data Science Mini-Symposium (SIAM Conference on Mathematics of Data Science).	2020
· IBM TJ Watson Research Center.	2019
· ICML Coding Theory for Large-scale Machine Learning Workshop (Spotlight).	2019
 Modeling and Optimization: Theory and Applications (MOPTA), Lehigh University. 	2018
· University of Helsinki, Finland.	2018
· IBM TJ Watson Research Center.	2017

TEACHING AND EDUCATIONAL CONTRIBUTIONS

• Foundations of Machine Learning (ENEE436) - Undergraduate	<i>E Level Course</i> Fall 2022, 2023, 2024
Instructor for this course at University of Maryland, College Park	ς.
• Estimation and Detection Theory (ENEE621) - Graduate Level	<i>Course</i> Spring 2023

Instructor for this course at University of Maryland, College Park

Educational contributions at CMU prior to starting a faculty position at UMD:

• New-Age Information Theory (18753) - *Graduate Level Course* Spring 2020 I contributed to the design and preparation of new course content, taught some lectures, and also prepared new assignments for the course. This course was a redesign of the previously offered course on information theory. Our redesign emphasized on both classical information and coding theory, as well as, emerging bodies of work within information theory, such as Partial Information Decomposition, and novel applications of information theory to machine learning, fairness, neuroscience, etc.

•	Information Theory (18753) - Graduate Level Course	Spring 2018
	Teaching Assistant	
•	Signals and Systems (18290) - Undergraduate Level Course Teaching Assistant	Fall 2017
•	Foundations of Cloud and ML Infrastructure (18847F) - <i>Graduate Level Course Guest Lecturer</i>	Fall 2017, 2018, 2019

• Foundations of Privacy (18734) - *Graduate Level Course* Fall 2020 Assisted in the design of course projects based on the CMU ECE graduate admissions dataset that I obtained after an Institutional Review Board (IRB) approval.

MENTORSHIP

PhD Students

- Faisal Adamu Hamman (Ph.D. student, University of Maryland College Park).
- Pasan Dissanayake (Ph.D. student, University of Maryland College Park).
- Barproda Halder (Ph.D. student, University of Maryland College Park).

Masters Students

- Allen Sunny (University of Maryland, iSchool).
- Pranjal Atrey (University of Maryland, ECE).
- Yiyue Zhu (University of Maryland, ECE).

Undergraduate Students

- Justin Cargiulo (University of Maryland, ECE).
- Kent Chelsey (University of Maryland, ECE).
- Aryan Kakadia (University of Maryland, ECE).
- Rohan Awasthi (University of Maryland, CS).

Student mentorship prior to starting a faculty position at UMD:

- Capstone Project (Collaboration between JP Morgan and Columbia University) Students: Xin Ye, Xiong Yue, Yuzhao Pan, Liyi Zhang, Panyu Gao (Spring 2022).
- Capstone Project (Collaboration between JP Morgan and Columbia University) Students: Junzhi Ge, Mohammed Aqid Khatkhatay, Oscar Jasklowski, Xue Gu, Yue Wang (Fall 2021).
- Utsav Sheth (INI, CMU), Fall 2017 2018.
- Malhar Chaudhuri (ECE, CMU), Fall 2016 2017.
- Chenyu Jiang (Visiting Undergraduate Student at CMU from CUHK), Summer 2019.
- Bowen Wu (Visiting Undergraduate Student at CMU from CUHK), Summer 2019.
- Ziqian Bai (Visiting Undergraduate Student at CMU from CUHK), Summer 2017.

PROFESSIONAL ACTIVITIES

Workshops and Panels

- Organizing Committee: UMD Machine Learning Rising Stars Workshop'23.
- Co-organizer: Multimodal AI for Financial Forecasting (Muffin) Workshop in AAAI'23.
- Co-organizer: XAI Workshop on Explainability in Finance at ICAIF'2023.
- Panelist: AutoML Conference AutoML and Fairness Session.

Membership in Professional Societies/Committees

- Member of the Organizing Committee of Energy & Information Systems (EIS) Seminar, ECE, CMU.
- Student Member of Diversity Committee, ECE, CMU.

Reviewing Responsibilities

- Area Chair, FAccT'24.
- Guest Editor, Entropy.
- Program Committee: Advances on Neural Information Processing Systems (NeurIPS), International Conference on Machine Learning (ICML), AAAI Conference on Artificial Intelligence (AAAI), International Conference on Artificial Intelligence and Statistics (AISTATS), ACM Conference on Fairness, Accountability, and Transparency (FAccT).
- Reviewer: IEEE Symposium on Information Theory (ISIT), IEEE Transactions on Information Theory, IEEE Transactions on Communications, IEEE Transactions on Signal Processing, IEEE Transactions on Signal and Information Processing over Networks, IEEE Journal on Selected Areas in Information Theory (Web of Science Reviewer Profile).

Outreach

- Co-Chair of Women In Information Theory Society (WITHITS).
- Faculty Research Mentor with Montgomery Blair High School (Magnet Program): Mentor high-school students on summer projects; students accrue course credits and communicate their findings by writ-

ing a scientific paper, preparing a presentation, and creating a poster at the Blair Research Convention.

• Faculty Lead at AI Summer Academy, Iribe Initiative for Inclusion and Diversity, UMD: Led experiential learning projects on Explainable AI at this summer camp targeted towards 10th through 12th graders; program aims to foster an inclusive and diverse field of AI by serving underrepresented communities.

POSITIONS OF RESPONSIBILITY

President, Indian Graduate Student Association (IGSA), CMU.	2018
Vice President, Indian Graduate Student Association (IGSA), CMU.	2017
General Secretary, Indian Graduate Student Association (IGSA), CMU.	2016
ECE Representative, Graduate Student Assembly (GSA), CMU.	2016-2019

EXTRA-CURRICULAR ACTIVITIES

Dramatics: Member of Druheen (formerly called Bengali Technology Drama Society) at IIT Kharagpur. **Fine Arts:** Qualification from Pracheen Kala Kendra, one of the oldest cultural organizations of India. **Community Service:** Volunteer of National Service Scheme (NSS), India. **Mentorship:** Student Mentor in collaboration with Student Welfare Group (SWG) at IIT Kharagpur.