

ARPIT MERCHANT

CONTACT	T0361, Morgan 2 Sanger Institute Wellcome Genome Campus Hinxton, CB10 1SA, UK Email : am84@sanger.ac.uk arpitdm@gmail.com Personal Webpage : https://www.arpitmerchant.com/
STATUS	U.K. Residency (with work permit), Finnish Residency (with work permit), Indian Citizen
RESEARCH INTERESTS	My research spans the areas of Machine Learning, Data Management, and Algorithmic Ethics in the context of responsible design of Graph-based AI systems. I develop models and methodologies to identify and enhance normative goals including robustness, scalability, and reliability in large-scale AI systems with particular applications to Spatial Transcriptomics and Single-Cell Biology.
EDUCATION	Doctor of Philosophy, Computer Science 2019 - 2023 University of Helsinki, Finland Thesis : Applications of Node Embeddings to Learning and Mining on Graphs Advisor : Michael Mathioudakis Bachelor of Technology (Honours), Computer Science 2011 - 2017 with Master of Science, Exact Humanities International Institute of Information Technology Hyderabad, India Thesis : The Use of Trust in Social Machines Advisor : Navjyoti Singh
PUBLICATIONS	Refereed Journal Publications [J1] Merchant A. , Mahadevan A., Mathioudakis M. <i>Scalably Using Node Attributes and Graph Structure for Node Classification</i> , Entropy, Volume 24, No. 7, 2022. Refereed Conference Proceedings [C10] Merchant A. , Castillo C. <i>Disparity, Inequality, and Accuracy Tradeoffs in Graph Neural Networks for Node Classification</i> in Proceedings of the 32nd International Conference on Information Knowledge and Management (CIKM), Birmingham, UK, October 21-25, 2023. [C9] Li, J., Wang, Y., Merchant A. . <i>Spectral Normalized-Cut Graph Partitioning with Fairness Constraints</i> in Proceedings of the 26th European Conference on Artificial Intelligence (ECAI), Kraków, Poland, September 30 - October 5, 2023. [C8] Merchant, A. , Mathioudakis, M., Wang, Y. <i>Graph Summarization via Node Grouping : A Spectral Algorithm</i> in Proceedings of the 16th ACM International Conference On Web Search And Data Mining (WSDM), Singapore, Singapore, February 27 - March 3, 2023. [C7] Mahadevan A.*, Merchant A.* , Wang Y., Mathioudakis M. <i>Robustness of Sketched Linear Classifiers to Adversarial Attacks</i> in Proceedings of the 31st International Conference on Information and Knowledge Management (CIKM), Atlanta, USA, October 17-21, 2022. [C6] Merchant A. , Gionis A, Mathioudakis M. <i>Succinct Graph Representations as Distance Oracles : An Experimental Evaluation</i> in Proceedings of the 48th International Conference on Very Large Databases (VLDB), Sydney, Australia, September 5-9, 2022.

- [C5] **Merchant A.**, Mathioudakis M. *Joint Use of Node Attributes and Proximity for Node Classification* in Proceedings of the International Conference on Complex Networks and Their Applications. (Complex Networks), Madrid, Spain, November 30 - December 2, 2021.
- [C4] **Merchant A.**, Shah D., Bhatia G. S., Ghosh A., Kumaraguru P. *Signals Matter : Understanding Popularity and Impact of Users on Stack Overflow* in Proceedings of the 30th ACM Conference on World Wide Web (WebConf), San Francisco, USA, May 13-17, 2019.
- [C3] Yeo T., Parameswaran K., Singla A., **Merchant A.**, Faucon P., Asselborn T., Dillenbourg P., Cevher V. *Iterative Classroom Teaching* in Proceedings of the 33rd AAAI Conference on Artificial Intelligence (AAAI), Hawaii, USA, January 27 - February 1, 2019.
- [C2] Tschitschek S., Singla A., Gomez-Rodriguez M., **Merchant A.**, Krause A. *Detecting Fake News in Social Networks via Crowdsourcing* in Proceedings of the 29th ACM Conference on World Wide Web (WebConf), Lyon, France, April 23-27, 2018.
- [C1] **Merchant A.**, Singh N. *Hybrid Trust-Aware Model for Personalized Top-N Recommendation* in Proceedings of the 4th ACM Conference of Data Science (CODS), Chennai, India, March 8-11, 2017. **Best Poster Award.**

Refereed Workshop Papers

- [W3] **Merchant A.**, Mathioudakis M. *Graph Summarization and Graph Embeddings : Towards A Spectral Connection* in the 16th International Workshop on Mining and Learning with Graphs at Knowledge Discovery and Data Mining (MLG @ KDD), Virtual, August 24, 2020.
- [W2] **Merchant A.**, Jha T., Singh N. *The Use of Trust in Social Machines* in the 4th International Workshop on the Theory and Practice of Social Machines at World Wide Web Conference, (Sociam @ WebConf), Montreal, Canada, April 11-15, 2016.
- [W1] **Merchant A.**, Shah D., Singh N. *In Wikipedia We Trust* in the 2nd Wiki Workshop at International Conference on Web and Social Media, ICWSM, Cologne, Germany, May 17-20, 2016.

Refereed Manuscripts/Reports

- [M3] **Merchant A.**, Jha T., Shukla M. *From Snail Post to Facebook Posts : A Case Study of Information Spread*. Mathematical Contest in Modelling (COMAP), Hyderabad, India, 2016. **Meritorious Award.**
- [M2] **Merchant A.**, Jha T., Shukla M. *Stochastic Distributed Decision Model for Combating Spread of Ebola*. Mathematical Contest in Modelling (COMAP), Hyderabad, India, 2015.
- [M1] **Merchant A.**, Prahladh H. *Multiplicative Weights Update : A Useful Addition to the Algorithmist's Toolkit*, Tata Institute for Fundamental Research (TIFR), Mumbai, India, 2013.

EXPERIENCE

Senior Data Scientist (postdoc) July 2024 - present
AI and Cell Engineering (with Mo Lotfollahi), Wellcome Sanger Institute, United Kingdom

- Curation and management of largest (to-date) spatial transcriptomics datasets of human cells.
- Foundational AI model for cell and gene level tasks.
- Niche-level generative modeling for estimating disease prioritization and progression.

Graduate Research Assistant January 2019 - October 2023
Algorithmic Data Science Group (with Michael Mathioudakis), University of Helsinki, Finland

- Developing theory, algorithms, and applications of low-dimensional representations of graphs, specifically focusing on node embeddings.
- Conducted analyses on node classification, summarization, and graph property estimation.
- Designed large-scale experiments on graphs with millions of nodes from social, communication, web, and biological network domains.

Visiting Doctoral Researcher July 2021 - January 2023
with Carlos Castillo, University Pompeu Fabra, Spain and University of Helsinki, Finland

- Quantifying the tradeoffs between algorithmic fairness and accuracy of graph neural networks in the context of node classification.
- Designed pre-training and post-training interventions to reduce algorithmic bias.

Research Associate

March 2018 - November 2018

PreCog Group (with Ponnurangam Kumaraguru), IIIT-Delhi

- Studied the role of game elements such as badges and reputation points in characterizing underlying social qualities like popularity and impact of users on Stack Overflow.
- Applied statistical learning, and time-series analysis methods to empirically quantify and validate the strength of game elements/digital signals on a rich dataset of 3,831,147 users and their activities spanning over a decade.
- Analyzed the presence of costly to earn and hard to observe signals to qualitatively differentiate between highly impactful and highly popular users.

Visiting Research Scholar

September 2017 - February 2019

Machine Teaching Group (with Adish Singla), MPI-SWS, Germany

- Designing novel, state-of-the-art algorithms for teaching a classroom of online projected gradient-descent learners with provable guarantees under complete and incomplete information paradigms.
- Studied applications of machine teaching methods on synthetic and real-world data for binary classification and handwriting improvement tasks.
- Developed Bayesian inference algorithms for detecting fake news in social networks and jointly learning users' flagging accuracy over time.

Junior Research Fellow

September 2016 - July 2017

Network Science Lab (with Anirban Dasgupta), IIT-Gandhinagar, Germany

- Designed versatile algorithms and novel, unbiased estimators for degree distribution and degree-wise clustering coefficients of large graphs.
- Conducted experimental evaluations on public network datasets to obtain close estimates of the actual values, for storage less than 1% of the input graph size.

Google Summer of Code Intern

May 2016 - August 2016

Sagemath (with Johan Rosenkilde and David Lucas)

- Designed computationally efficient methods for atomic representations of elements of skew polynomial rings and for answering basic questions about them.
- Implemented methods for computing combinatorial properties, and for encoding and decoding algorithms for Golay, Gabidulin and Rank-Metric Codes.

Research Intern

May 2013 - July 2013

Tata Institute for Fundamental Research (with Prahladh Harsha)

- Analyzed randomized variants of the deterministic multiple weights update algorithm to show that the total expected cost is only slightly worse than the best possible strategy in hindsight.
- Surveyed its application to the multicommodity packing flows problem and a high quality approximation efficiently with provable guarantees.

TEACHING & SERVICE

Program Committee : WSDM, DSAA

Reviewing : TWEB, ICDE, KDD, VLDB

Teaching Assistant : Graduate course on Network Analysis. Undergraduate courses on Introduction to Discrete Mathematics, Introduction to Probability, Optimization Methods. Responsibilities included designing homework assignments, guiding student projects, lecturing, and grading.

Teaching Fellow : Computational Thinking and Applications for Middle and High School Stu-

dents. Responsibilities included designing curriculum, lecturing, conducting lab sessions.

Science Communication : Skype A Scientist, Theory Reading Group

- INVITED TALKS
- Graph Summarization via Node Grouping, CLAWS Group Seminar, Georgia Institute of Technology, USA, October, 2022.
 - Jointly Using Node Attributes and Graph Structure for Classification, Complex Systems Seminar, Aalto University, Finland, September, 2022.
 - Merchant A., Choudhary J., Dasgupta A., Efficiently Estimating Primitive Graph Properties, Max Planck Institute for Informatics, Saarbruecken, Germany. June 2017.
 - Merchant A., Rosenkilde J., Lucas D. Rank-Metric Codes in SageMath, Sage Days 75, INRIA Ile-de-Saclay, Paris, France. August, 2016
 - Merchant A., Singh N. The Use of Trust in Social Machines, IIIT-H Spring Research and Development Showcase, Hyderabad, India. March, 2016.

SKILLS

Programming : Python, C++, C, SQL, Bash, Git

Libraries/Tools : PyTorch, Pytorch-Geometric, NetworkX, Pandas, Scikit-Learn

- AWARDS & HONORS
- Doctoral Grant, Helsinki Institute of Information Technology, 2023.
 - DONASCI Travel grant, University of Helsinki, Finland, 2023.
 - Nokia PhD Fellowship, 2022.
 - Student Travel grant, ACM IKDD Conference of Data Science (CODS), Indian Institute of Technology, Chennai, 2017.
 - Awarded Dean's Merit List (top 5% of the class) for Academic Excellence, 2014.
 - Special Commendation Award (second-highest student award at IIIT Hyderabad) for "...exceptional inspirational contributions and for upholding and highlighting the importance of being responsible.", 2015.
 - National Research Fellowship, jointly sponsored by IASc, INSA, NASI, India, 2013.

- REFERENCES
- Dr. Michael Mathioudakis**
Associate Professor
University of Helsinki, Finland
Email : michael.mathioudakis@helsinki.fi
Webpage : <https://michalis.co/>
- Dr. Carlos Castillo**
ICREA Research Professor
Universitat Pompeu Fabra, Spain
Email : chato@icrea.cat
Webpage : <https://chato.cl/>
- Dr. Mo Lotfollahi**
Group Leader, Faculty
Wellcome Sanger Institute, United Kingdom
Email : m119@sanger.ac.uk
Webpage : <https://lotfollahi.com>
- Dr. Aristides Gionis**
WASP Professor
KTH Royal Institute of Technology, Sweden
Email : argioni@kth.se
Webpage : <https://www.kth.se/profile/argioni>