Madhurima Nath

madhurimanath21@gmail.com LinkedIn: madhurimanath | GitHub: mnathvt

Multi-cloud certified machine learning scientist and data engineer, proficient in algorithm development and computational modeling.

Education

Ph.D., Physics, Virginia Tech, Blacksburg, VA Dec 2018
Dissertation: *Application of Network Reliability to Analyze Diffusive Processes on Graph Dynamical Systems*

This work explores the effects of the structural properties of an interacting system on the outcomes of a diffusive process on realistic socio-technical systems using an efficient and generalized probabilistic measure based on Monte-Carlo simulations and graph theory techniques.

US Patent: *System, method and computer readable medium for sensitivity of dynamical systems to interaction network topology*

M.S., Physics, Virginia Tech, Blacksburg, VA May 2017

M.Sc., Physics, Indian Institute of Technology Delhi, New Delhi, India May 2012
Thesis: *Study of Cold Atomic Condensates by Atomic Photon Interactions*
Award: *Best Master of Science Thesis 2011-2012*

B.Sc. (Hons.), Physics, University of Calcutta, Kolkata, India May 2010
Minors: *Mathematics and Chemistry*

Skills

Programming: Python, SQL, PySpark, R || **CI/CD**[1]: Git/GitHub/Azure DevOps/BitBucket
Tools: Databricks, Azure Machine Learning (ML), Tableau, AWS[2] QuickSight, Power BI, Mathematica, MATLAB, Shell scripting
Certifications: Azure Data Scientist Associate, Databricks Lakehouse Data Engineer Associate, GCP[3] Associate Cloud Engineer, AWS[2] Cloud Practitioner, Domino Data Lab Data Science Practitioner, Tableau Desktop Specialist, Matillion Associate
Other Skills: Data Modeling, Gap Analysis, Technical, Conceptual and Data Architecture Diagrams, Data Flows, Agile Delivery
Awards: Slalom Mogul Award recognizing outstanding client service and delivery excellence, Q1 2023, Q2 2022, Q4 2020.

Industry Experience

Data Science Consultant, Data & Analytics, Slalom Consulting, LLC, New York, NY **Jan 2020 – Present**

- Co-led the development and deployment of a text classification model using NLP[4] techniques and XGBoost in Azure to categorise product data into multiple groups defined by a standardised template. This accelerated the data foundation launch roadmap by over a year and provided a comprehensive view of $35B global procurement spend for a consumer packaged goods client.
- Led the development and deployment of sentiment analysis and topic modeling NLP[4] pipelines in Azure Databricks to uncover safety-related themes from survey-data auto-refreshed weekly. The executive safety committee of a large Midwest utility provider used the resulting PowerBI dashboard to identify risky practices and improve the policies to safeguard their employees.
- Designed and implemented a prototype interactive knowledge-based tool to assist sales representatives to service customer inquiries in an efficient and concise manner, using AWS[2] managed services, as the first digital innovation project for one of the largest equipment rental clients.
- Co-led the design and development of fuzzy matching rules and data engineering pipelines to integrate syndicated retail data from multiple vendors - including Nielson, IRI and Skupos - into a centralized platform on Azure, replacing existing manual processes and improving efficiency for product comparison by $\sim 28\%$ for a consumer packaged goods client.
- Co-led multiple solution accelerators initiatives on Azure Databricks, including data ingestion pipelines using medallion structure, data quality frameworks, NLP[4] solution to implement semantic search, to jumpstart product delivery for clients.
- Coached and mentored summer interns, served as go-to resource on advanced analytics solutions and computational modeling.
- Created a segmentation model using historical customer viewership of sports events to implement better reporting, personalized marketing strategies and data governance for a media and television client.
- Built a proof-of-concept solution to evaluate propensity scores and feature importance for accounts based on historical data in Salesforce, enabling sales & accounts teams to identify and prioritize potential buyers with $\sim 93\%$ accuracy, replacing manual efforts.
- Built Databricks pipelines to process data from medallion framework and AWS[2] QuickSight dashboards to provide quick insights into patient data related to treatments and clinical trials for a multinational biotechnology client.
- Performed gap analysis, gathered business requirements and analyzed customer engagement from historical data in Snowflake to define key performance metrics of a new product launch for a media and television client.
- Built scalable interactive visualization dashboards in Tableau for sales and commissions related analytics, enhancing existing solutions for a telecom client.

[1] *CI/CD: Continuous Integration & Continuous Development*, [2] *AWS: Amazon Web Services*, [3] *GCP: Google Cloud Platform*, [4] *NLP: Natural Language Processing*

Research Experience

Post-graduate Research Assistant, Virginia Tech, VA **Mar 2019 – Dec 2019**

- Developed an algorithm using Monte-Carlo simulations and perturbative methods to efficiently estimate the solution of a NP-hard problem with sufficiently good precision for all practical purposes.
- Applied to food networks data from United Nations Comtrade database to detect significant communities important to prevent global distribution of pestilence, this method outperforms the well-known heuristic solution by $\sim 12\%$ on average.

Graduate Research Assistant, Virginia Tech, VA **May 2014 – Dec 2018**

- Developed an algorithm to efficiently approximate real-world scenarios for epidemic outbreaks as a function of interactions and infection rate. Policymakers can be equipped with better and more accurate estimates, instead of using existing methods which overestimate the number of infections by $\sim 50\%$.
- Identified vulnerabilities in international food trade networks and estimated effects of mitigating contagion conditions with 96% accuracy to guide informed decision-making.
- Developed an algorithm using concepts of Moore–Shannon network reliability to estimate the energy states of an interacting magnetic system, reducing the time complexity from $O(n^2)$ to $O(n)$.

Instructor of Record, Department of Physics, Virginia Tech, VA **Aug 2018 – Dec 2018**

- Ranked as one of the most effective physics instructors by students who fosters an atmosphere of mutual respect.
- Designed and facilitated weekly lectures, one-on-one sessions and recitation sessions for a 3-credit introductory physics course curriculum for ~ 120 undergraduate engineering students.
- Supervised physics laboratory sessions and office hours for both engineering and non-engineering students, which improved the grades of $\sim 15\%$ of the students.

Publications

- S. Eubank, **M. Nath**, R. Mishra, and A. Adiga, Communities in directed weighted food networks using Moore-Shannon network reliability. (submitted Applied Network Science).
- R. Mishra, S. Eubank, **M. Nath**, M. Amundsen and A. Adiga, Community detection using Moore-Shannon network reliability: Application to food networks, Complex Networks and Their Applications XI: Volume 2 Proceedings of The 11th International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2022, Springer Insternational Publishing, 271-282, 2023.
- S. Eubank, **M. Nath**, Y. Ren and A. Adiga, Perturbative methods for mostly monotonic probabilistic satisfiability problems (arXiv preprint arXiv:2206.0355).
- **M. Nath** et. al., Using network reliability to understand international food trade dynamics, Complex Networks and Their Applications VII: Volume 1 Proceedings The 7th International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2018, Springer International Publishing, 524- 535, 2019.
- **M. Nath**, Y. Ren and S. Eubank, An approach to structural analysis using Moore-Shannon network reliability, Complex Networks and Their Applications VII: Volume 1 Proceedings The 7th International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2018, Springer International Publishing, 537-549, 2019.
- **M. Nath**, Y. Ren, Y. Khorramzadeh, and S. Eubank, Determining whether a class of random graphs is consistent with an observed contact network, J. Theor. Biol. 440C, 121-132, 2018.
- **M. Nath** and S. Eubank, Model selection for sequential designs in discrete finite systems using Bernstein kernels (arXiv preprint arXiv:1807.06661).
- Y. Ren, S. Eubank, and **M. Nath**, From network reliability to the Ising model: A parallel scheme for estimating the joint density of states, Phys. Rev. E 94.4: 042125, 2016.
- **M. Nath**, S. Eubank, M. Youssef, Y. Khorramzadeh, and S. Mowlaei, A two-parameter method to characterize the network reliability for diffusive processes, Complex Networks VI: Proceedings of the 6th Workshop on Complex Networks CompleNet 2015, Springer International Publishing 139-148, 2015.
- A. Agarwala, **M. Nath**, J. Lugani, K Thyagarajan, and S. Ghosh, Fock-space exploration by angle resolved transmission through a quantum diffraction grating of cold atoms in an optical lattice, Phys. Rev. A 85.6: 063606, 2012 (*equal contribution as first author*).

Conference Presentations

- **Invited Talk**: Network reliability: a generic tool to explore diffusive processes on interacting systems, NASA PCE3 (Prebiotic Chemistry and Early Earth Environments) Virtual Workshop 2022 Nano- to Cosmic- Studies of Complex Systems, University of Wisconsin Madison, Madison, WI, Oct. 20, 2022.
- **Talk**: Perturbative methods for estimating relative contributions to network reliability, SIAM (Society for Industrial and Applied Mathematics) Workshop on Network Science, Virtual Workshop, Sep 13-15, 2022.
- **Talk**: Statistical mechanical applications of graph dynamical systems, Condensed Matter Seminar, Department of Physics, Virginia Tech, Blacksburg, VA, Oct. 30, 2017.
- **Talk**: Determining whether a particular contact network is consistent with a network model, 1st North American Social Networks Conference of the International Network for Social Network Analysis, Washington DC, Jul. 26-30, 2017.
- **Talk**: Network reliability: A novel measure to study the effects of network topology on the diffusive dynamics, Symposium for the Society of Young Network Scientists, NetSci 2017, Indianapolis, IN, Jun. 19-23, 2017.
- **Poster**: Network reliability: A measure to study diffusive dynamics on networks, Center for Soft Matter and Biological Physics Symposium 2017, Virginia Tech, Blacksburg, VA, May 17-18, 2017.
- **Talk**: Effects of network structure on propagation of infectious diseases, 33rd Annual Graduate Student Assembly Symposium and Exposition, Virginia Tech, Blacksburg, VA, Mar. 29, 2017.
- **Talk**: Renormalization group approaches for dynamics on irregular networks, APS (American Physical Society) March Meeting 2017, New Orleans, LA, Mar. 13-17, 2017.
- **Poster**: Diffusive dynamics on a network, SESAPS (Southeastern Section of the American Physical Society) Conference 2016, Charlottesville, VA, Nov. 9-12, 2016.
- **Poster**: Effects of network structure on epidemic modeling, Biocomplexity Institute Symposium 2016, Virginia Tech, Blacksburg, VA, Nov. 1, 2016.
- **Talk**: A two-parameter method to characterize the network reliability for diffusive processes, CompleNet 2015, New York City, NY, Mar. 25-27, 2015.
- **Poster**: Four-parameter characterization of network reliability and analysis of critical point phenomenology, APS (American Physical Society) March Meeting 2015, San Antonio, TX, Mar. 2-6, 2015.

Outreach

- Reviewer - Synergy of Scientific and Machine Learning Modeling, 2023 International Conference on Machine Learning. 2023
- Reviewer Machine Learning and the Physical Sciences, NeurIPS2022 - Conference on Neural Information Processing Systems. 2022
- Reviewer - Journal - Physical Review E. 2022 – Present
- Industry Ambassador, Women in Network Science Society. Sep 2022 – Present
- Data science instructor at Slalom Q2 2022 New York Learning Session. Jun 2022
- Speaker at Women Who Code Data Science and San Francisco Backend chapters. 2Apr – Jul 2021
- Invited Speaker and Panelist at Women Who Code Statistics in Data Science Workshop Series Feb – Mar 2021
- Organizer of Women in Network Science Networks 2021 Conference Jun – Jul 2021
- Invited student speaker of APS Conference for Undergraduate Women in Physics, 2017. Jan 2017

References

- Prof. Stephen Eubank
  Professor, Biocomplexity Institute & Innovation, University of Virginia
  email: eubank@virginia.edu
- Prof. Madhav Marathe
  Professor, Biocomplexity Institute & Innovation, University of Virginia
  email: marathe@virginia.edu
- Prof. Uwe Täuber
  Professor, Department of Physics, Virginia Tech
  email: tauber@vt.edu
- Maria Mendoza
  Senior Data Engineer, Gemini
  email: auimendoza@gmail.com
- Prof. Lara B. Anderson
  Professor, Department of Physics, Virginia Tech
  email: lara.anderson@vt.edu