# Mononito Goswami

Newell-Simon Hall, Carnegie Mellon University
Pittsburgh, PA-15213, USA
mgoswami [at] andrew [dot] cmu [dot] edu
LinkedIn | ResearchGate | Google Scholar | Website

RESEARCH INTERESTS

Foundation Modeling, Modeling Structured Data (time series and tabular), Machine Learning (ML), ML for Healthcare, Weak Supervision

EDUCATION

*Doctor of Philosophy in Robotics* 2020 - 2025 (expected)
Carnegie Mellon University, Pittsburgh PA, USA

- Doctoral Dissertation: Towards Pragmatic Time Series Intelligence
- Advisor: Prof. Artur Dubrawski

*Bachelor of Technology in Computer Engineering* 2016 - 2020
Delhi Technological University, New Delhi, India

- Thesis: Towards Social & Engaging Peer Learning [Paper 1, Paper 2]

FELLOWSHIPS

Centre for Machine Learning and Health (CMLH) 2021 2021 - 2022

RESEARCH EXPERIENCE

*Student Researcher* May 2023 - Present
Google Research, New York, USA

- Machine Learning research on building, pre-training and evaluating **foundation models for tabular data**, along with Scott Yak, Joe Toth, Yihe Dong, Sercan Arik.

*Applied Scientist Intern* May - August 2023
Amazon Web Services AI Labs, Seattle, USA

- Ideated and built one of the first **time series foundation models**, along with Bariş Kurt, Andrey Kan, Gauthier Guinet, Jingchao Ni, Jonas Kübler and Laurent Callot.

*Applied Scientist Intern* May - August 2022
Amazon Web Services AI Labs, Seattle, USA

- Developed one of the first algorithms for unsupervised model selection of time series anomaly detection models, in collaboration with Andrey Kan, Lenon Minorics and Laurent Callot [Paper].

*Robotics Institute Summer Scholar* June 2019 - August 2020
Auton Lab, Carnegie Mellon University, Pittsburgh, USA

- Machine Learning research on detecting cognitive disequilibrium and flow in children solving math problems, advised by Prof. Lujie (Karen) Chen and Prof. Artur Dubrawski [Paper, Student abstract].

*Robotics Institute Summer Scholar* June 2018 - September 2020
RoboTutor Project, Carnegie Mellon University, Pittsburgh, USA

- Developed Statistical Probe of Tutoring (SPOT), a tool for iterative data-driven improvement of RoboTutor, an Intelligent Tutoring System (ITS), advised by Prof. Jack Mostow [Paper, Student abstract].

*Undergraduate Researcher* 2017 - 2020
Delhi Technological University, New Delhi, India

- Analyzing dyadic interactions between young children to identify non-verbal cues that aid effective story-telling, advised by Prof. Rajni Jindal [Paper 1, Paper 2].
- Developed a Multi-task Learning approach for Open Domain Suggestion Mining and a novel language model-based text over-sampling method, advised by Ms. Minni Jain [Paper, Student abstract].
- Improvised energy-efficient clustering & routing algorithms for Wireless Sensor Networks using modified Binary Particle Swarm Optimization, advised by Prof. Indu S and Prof. Daya Gupta [Paper].
- Designed an Intrusion detection algorithm for critical RBAC administered databases using Pattern Mining and nearest-neighbours Anomaly Detection, advised by Ms. Indu Singh [Paper].
- Investigating applications & modelling of fractional order-differential equations (FODEs) for control of infectious diseases using SVEIR models, advised by Dr. Nilam [Report].
- Distracted driver detection in real-time using a simple CNN-model. Advisors: Dr. Rajiv Ratn Shah, Dr. Yifang Yin and Dr. Roger Zimmermann [Paper].

GUEST LECTURE (at CMU)

- Implicit Communication and Theory of Mind (for 16-467– Human-Robot Interaction)

TEACHING ASSISTANTSHIP (at CMU)

- 16-811 – Math Fundamental for Robotics — Fall 2022
- 16-467 – Human-Robot Interaction — Spring 2022

CONFERENCE & JOURNAL ARTICLES

See also my google scholar page. ** indicates equal contribution*

15. **Goswami, Mononito**, Konrad Szafer*, Arjun Choudhry*, Yifu Cai, Shuo Li, and Artur Dubrawski. "*MOMENT: A Family of Open Time-series Foundation Models.*" *International Conference on Machine Learning (ICML 2024).* [PDF, Code, Pre-trained Model, Pre-training Dataset]
14. **Goswami, Mononito**, Vedant Sanil, Arjun Choudhry, Arvind Srinivasan, Chalisa Udompanyawit, Artur Dubrawski. "*AQuA: A Benchmarking Tool for Label Quality Assessment.*" *Neural Information Processing Systems (NeurIPS 2023) Track on Datasets and Benchmarks.* [PDF, Code] (**Poster**)
13. **Goswami, Mononito**, Cristian Challu, Laurent Callot, Lenon Minorics, and Andrey Kan. "*Unsupervised Model Selection for Time-series Anomaly Detection.*" *International Conference of Learning Representations (ICLR 2022).* [PDF, Code] (**Spotlight**)
12. Gao, Chufan*, **Mononito Goswami***, Jieshi Chen and Artur Dubrawski. "*Classifying Unstructured Clinical Notes via Automatic Weak Supervision.*" *Machine Learning for Healthcare Conference (MLHC 2022).* [PDF, Code]

11. Dey, Arnab, **Mononito Goswami**, Joo Heung Yoon, Gilles Clermont, Michael R. Pinsky, Marilyn Hravnak, Artur Dubrawski. "*Weakly Supervised Classification of Vital Sign Alerts as Real or Artifact.*" In *AMIA Annual Symposium Proceedings*. American Medical Informatics Association. [PDF, Code]
10. Nagpal, Chirag, **Mononito Goswami**, Keith Dufendach, and Artur Dubrawski. "*Counterfactual Phenotyping with Censored Time-to-Events*". (2022) In *ACM Conference on Knowledge Discovery and Data Mining*. [PDF, Code]
9. **Goswami, Mononito**, Benedikt Boecking, and Artur Dubrawski. "*Weak Supervision for Affordable Modeling of ECG Data.*". (2021) In *AMIA Annual Symposium Proceedings*. American Medical Informatics Association. [PDF]
8. McReynolds, Andrew A., Sheba P. Naderzad, **Mononito Goswami**, and Jack Mostow. "*Toward Learning at Scale in Developing Countries: Lessons from the Global Learning XPRIZE Field Study.*" In Proceedings of the Seventh ACM Conference on Learning@ Scale, pp. 175-183. 2020. [PDF]
7. Singh, Indu, Minkush Manuja*, Rishabh Mathur*, and **Mononito Goswami***. "*Detecting intrusive transactions in databases using partially-ordered sequential rule mining and fractional-distance based anomaly detection.*" International Journal of Intelligent Engineering Informatics 8, no. 2 (2020): 138-171. [PDF].
6. Kaushik, Ajay*, **Mononito Goswami***, Minkush Manuja*, Indu S. and Daya Gupta. "*A Binary PSO Approach for Improving the Performance of Wireless Sensor Networks.*" Wireless Personal Communications (2020): 1-35. [PDF]
5. Jindal, Rajni*, Maitree Leekha*, Minkush Manuja*, and **Mononito Goswami***. "*What makes a better companion? towards social & engaging peer learning.*" In ECAI 2020, pp. 482-489. IOS Press, 2020. [PDF]
4. Leekha, Maitree*, **Mononito Goswami*** and Minni Jain "*A Multi-task Approach to Open Domain Suggestion Mining using Language Model for Text Over-sampling*". In: Jose J. et al. (eds) Advances in Information Retrieval. ECIR 2020. Lecture Notes in Computer Science, vol 12036. Springer, Cham [PDF]
3. **Goswami, Mononito***, Lujie Chen* and Artur Dubrawski. "*Discriminating Cognitive Disequilibrium and Flow in Problem Solving: A Semi-supervised Approach Using Involuntary Dynamic Behavioral Signals*". Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 34. 2020. [PDF]
2. Leekha, Maitree*, **Mononito Goswami***, Rajiv Ratn Shah, Yifang Yin and Roger Zimmermann. "*Are You Paying Attention? Detecting Distracted Driving in Real-time*". Proceedings of the IEEE International Conference on Multimedia Big Data (BigMM) [PDF]
1. Mian, Shiven*, **Mononito Goswami***, and Jack Mostow. "*What's Most Broken? Design and Evaluation of a Tool to Guide Improvement of an Intelligent Tutor.*" International Conference on Artificial Intelligence in Education. Springer, Cham, 2019 [PDF]

PEER-REVIEWED WORKSHOP PUBLICATIONS & ABSTRACTS

10. Choudhry, Arjun, Konrad Szafer, **Mononito Goswami**, Yifu Cai, and Artur Dubrawski. "*Datasets for Time Series Foundation Models*" ICML 2024 Workshop on Data-Centric Machine Learning Research (DMLR 2024). 2024. [PDF]
9. Cai, Yifu, Arvind Srinivasan, **Mononito Goswami**, Arjun Choudhry, and Artur Dubrawski. "*JoLT: Jointly Learned Representations of Language and Time-Series for Clinical Time-series Interpretation*" Proceedings of the AAAI Conference on Artificial Intelligence (Student Abstract). 2024. **Best student abstract presentation award winner.** [PDF]

8. Enouen, Eric, Sebastian Caldas, **Mononito Goswami**, and Artur Dubrawski. "*PICSR: Prototype-Informed Cross-Silo Router for Federated Learning*" Proceedings of the AAAI Conference on Artificial Intelligence (Student Abstract). 2024. *3-min presentation contest finalist.*
7. Cai, Yifu, **Mononito Goswami**, Arjun Choudhry, Arvind Srinivasan and Artur Dubrawski. "*JoLT: Jointly Learned Representations of Language and Time-Series.*" Neural Information Processing Systems Workshop on Deep Generative Models for Health (DGM4H NeurIPS) (2023) (Poster).
6. Caldas, Sebastian, **Mononito Goswami** and Artur Dubrawski. "*Encoding Expert Knowledge into Federated Learning Using Weak Supervision.*" International Conference of Learning Representations Workshop on Machine Learning for IoT (ICLR ML4IoT) (2023).
5. Rooney, Sydney R, Roman Kaufman, **Mononito Goswami**, Michael R Pinsky, J. Kyle Miller, Salah Al-Zaiti, Artur Dubrawski and Gilles Clermont. "*Using Weakly Supervised Machine Learning to Label Atrial Fibrillation in Real-World Intensive Care Unit Telemetry Data.*" Circulation 146.Suppl_1 (2022): A10198-A10198.
4. **Goswami, Mononito***, Lujie Chen*, Chufan Gao and Artur Dubrawski. "*Modeling Involuntary Dynamic Behaviors to Support Intelligent Tutoring (Student Abstract)*". Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 34. 2020. [PDF]
3. Gao, Chufan, Fabian Falck, **Mononito Goswami**, Michael R. Pinsky, Anthony Wertz and Artur Dubrawski. "*Detecting Patterns of Physiological Response to Hemodynamic Stress via Deep Unsupervised Learning*". Machine Learning for Health (ML4H) Workshop at NeurIPS 2019 [PDF]
2. Jain, Minni*, Maitree Leekha*, **Mononito Goswami***. "*A Multi-task Approach to Open Domain Suggestion Mining (Student Abstract)*". Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 34. 2020. [PDF]
1. **Goswami, Mononito***, Shiven Mian*, and Jack Mostow. ''*What's Most Broken? A Tool to Assist Data-Driven Iterative Improvement of an Intelligent Tutoring System.*" Proceedings of the AAAI Conference on Artificial Intelligence (Student Abstract). Vol. 33. 2019. *3-min presentation contest finalist* [PDF]

## INVITED TALKS

1. Time series Foundation Models– Challenges, Approaches, and Opportunities
   - AI & ML at CapitalOne — November, 2024
   - Datadog — November, 2024
   - Salesforce Research Asia — November, 2024
   - Commonwealth Bank of Australia — October, 2024
   - US Naval Center Warfare Center - Carderock Division — September, 2024
   - Forecasting Impact, Podcast by International Symposium of Forecasting — July, 2024
   - Prof. Xiao Hu's Lab, Emory University — April, 2024
   - CMU Flame Seminar — April, 2024
   - Gradient AI [webinar] — April, 2024
   - AAAI 2024 Spring Symposium on Clinical Foundation Models — March, 2024

MENTORSHIP (at CMU)

10. Nina Żukowska, Robotics Institute Summer Scholar Program (RISS), *now* Master's student at FU Berlin 2024–Present
9. Michał Wiliński, RISS 2024–Present
8. Konrad Szafer, RISS 2023–2024
7. Arjun Choudhry 2023 - Present
6. Yifu Cai 2023 - Present
5. Eric Enouen, *now* Ph.D. student at Cornell 2023
4. Undergrad AI Mentoring Program 2021 - 2022
3. Chalisa Udompanyawit, CIT Honors Research Program 2022 - 2023
2. Arnab Dey, RISS, 2021 - 2022
1. Graduate Application Support Program 2020

COMMITTEE MEMBERSHIP (at CMU)

4. Angela H. Chen, Ph.D. RI
3. Willa Potosnak, Ph.D. RI
2. Xinyu (Rachel) Li, Ph.D. RI
1. Ambareesh Revanur, Masters RI, *now* Adobe MLE

WORK EXPERIENCE

*Equity Research Intern* December 2017
Phillip Capital, Mumbai, India

- Carried out a study on disruptive technology like Blockchain & edge-computing that can potentially transform the *FinTech* sector.

*Intern* June - July 2017
Goods & Services Tax Network (GSTN), New Delhi, India

- Designed the Analytics & Risk Management framework along with consultants from PwC, Infosys and State Tax departments. Co-developed a simplified tool for tax submissions for the pan-India GST roll out.

*Intern* December 2016
Centre for Development in Advanced Computing (CDAC), Noida, India

- Developed a Grade-1 Unified English Braille (UEB) Conversion utility in C++. This work helped would help in implementation of UEB in India.

ACHIEVEMENTS

- Secured 1.2 million in research funding through leading and contributing to successful NSF grant proposals, including SCH: Multimodal Interactive Generalist Health AI (MAGENTA) and ATD: Spatiotemporal Foundation Models for Multimodal Threat Detection at Scale.
- Awarded for the best essay on *Goods & Services Tax*, its financial and technological implications, in the 2017 Indian Institute of Public Administration Essay Competition, by the *Vice President of India*.
- Secured second place among 500+ teams from colleges and startups across India in a National Payments Council of India (NCPI) Hackathon. Designed an intrusion detection architecture using fuzzy logic and keystroke dynamics.

SCHOLARSHIPS

- **Conference Travel Grants:** AAAI 2024, NeurIPS 2023, Microsoft Research Travel Grant for AAAI-20, AAAI-20 Student Scholarship, National Science Foundation Student Travel Grant to attend AIED 2019

PROFESSIONAL SERVICE

*Organization*

- Co-chaired and co-organized the highly successful [AAAI 2024 Spring Symposium on Clinical Foundation Models], attracting significant attendance within the AAAI Spring Symposium Series.

*Reviewer*

- AAAI– 2020
- American Medical Informatics Association (AMIA) 2021 Annual Symposium
- Journal of Electrocardiology
- ICLR– 2025, 2024, 2023, 2022
- ICML– 2024, 2023, 2021 ML4data workshop
- MLHC 2024
- NeurIPS– 2024 Main Track (Top Reviewer) & TSALM workshop, 2023 (Top Reviewer), 2022, 2021

*Admissions Committee*

- Robotics Institute Summer Scholar (RISS) - 2020, 2021, 2022

SOCIAL OUTREACH

- As a member of the Robotics Institute Climate Committee, identified challenges in the experiences of various groups within RI and made policy recommendations to the Director to address them. 2021 - Present
- Mentored two undergraduate students of an underrepresented groups interested in pursuing AI research, under the CMU AI Mentoring Program. 2020 - Present
- Exposed our research on RoboTutor to primary stakeholders, some 8-10 year olds from Pittsburgh schools and obtained interesting feedback for comparative cognitive processes, as a part of the Gelfand Outreach program. July 2018

PROGRAMMING Python, PyTorch, Jax

HOBBIES Chess, Photography, Cooking