SANA TONEKABONI

Postdoctoral Fellow, Broad Institute of MIT and Harvard
http://www.cs.toronto.edu/~stonekaboni/ stonekab@mit.edu

EDUCATION

Doctorate of Philosophy in Computer Science *2023*
University of Toronto - Advisor: Dr. Anna Goldenberg

Bachelor of Electrical Engineering with honors *2017*
University of Toronto (Minor in Bioengineering)

RESEARCH INTERESTS

Machine learning, Computational medicine, Time-series modelling and analysis, Self-supervised representation learning, Multimodal learning, Bayesian modelling, Uncertainty, Explainability in AI, Signal processing

HONORS AND AWARDS

Rising Star in Computational and Data Sciences *April 2023*

Apple's scholars in AI/ML Fellowship (USD $90K) *September 2021*

Canada Doctoral Graduate Scholarships (CAD $80K) *May 2021*
Natural Sciences and Engineering Research Council of Canada (NSERC)

Health System Impact Fellowship (CAD $100K) *September 2019*
Canadian Institute of Health Research (CIHR)

Queen Elizabeth II graduate scholarships in science and technology (CAD $5K) *September 2019*

General Motors women in electrical and mechanical engineering award *July 2015 and 2016*

NSERC undergraduate student research award (USRA) *May 2015 and 2016*

WORK EXPERIENCE

Postdoctoral Fellow *2024 - present*
Broad institute of MIT and Harvard

Research Intern - Apple *Summer 2022*
Apple health research team

Research Intern - Google *Summer 2021*
Google research cloud AI team

Applied Research Scientist Intern - Element AI *Fall 2018 - Summer 2019*
Uncertainty-aware classification models

Undergraduate Research Assistant - University of Toronto
Wireless communication lab - Advisor: Dr. Elvino Sousa *Summer 2016*
Intelligent sensory microsystem lab - Advisor: Dr. Roman Genov *Summer 2015*

TEACHING EXPERIENCE

Lecturer (Basic principles of machine learning in biomedical research-LMP1210) *Spring 2024*
University of Toronto [webpage]

Lecturer (Full-day practical session on machine learning for time series in health) *July 2023*
AI4Health Summer School (Paris, France)

Teaching Assistant - Tutorial (Statistical methods for machine learning) *Spring 2022*
University of Toronto

Teaching Assistant - Tutorial/Grading (Algorithms - Machine learning) *Fall 2019 - 2020*
University of Toronto

Teaching Assistant - Tutorial/Course development (Computing for medicine) *Fall - Spring 2020*
University of Toronto

PUBLICATIONS

Publications under review:

[C18] **S. Tonekaboni**, K. Alhamoud, M. Ghassemi. *Measuring Representation Uncertainty in Self-Supervised Embedding Models*. under review at CVPR 2026.

[C17] **S. Tonekaboni**, V. Schuster, C. Uhler. *Decoupling Shared and Modality-Specific Subspaces in Multimodal Learning via Low-Rank Representation Fine-Tuning*. under review at ICLR 2026.

[C16] V. Schuster, **S. Tonekaboni**, C. Uhler. *FiGuRO: Intrinsic Dimension Estimation for Multi-Modal Data*. under review at ICLR 2026.

[C15] Y. Xiao, **S. Tonekaboni**, W. Gerych, V. Suriyakumar, M. Ghassemi. *When Style Breaks Safety: Defending Language Models Against Superficial Style Alignment*. under review at ICLR 2026.

Published journal and conference papers:

[J4] **S. Tonekaboni**, S. L. Friedman, X. Zhang, M. Maddah, C. Uhler. (2025).
A Multimodal Representation Learning Framework to Learn a Unified Representation of Patients' Health with Dissentangled Modality-specific and Shared Information . accepted to npj Digital Medicine.

[C14] S. L. Friedman, **S. Tonekaboni**, A. Nargesi, C. Uhler, M. Maddah. (2025).
Cross-Attention Enables Phenotypic and Cross-Modal Conditioning of Biomedical Diffusion Models. In Machine Learning for Health (ML4H 2025).

[C13] **S. Tonekaboni**, L. Stempfle, A. Fallahpour, W. Gerych, M. Ghassemi. (2025). *An Investigation of Memorization Risk in Healthcare Foundation Models*. Advances in Neural Information Processing Systems, 33. presented at the Conference on Neural Information Processing Systems (NeurIPS 2025).

[C12] **S. Tonekaboni**, T. Behrouzi, A. Weatherhead, D. Blei, E. Fox, A. Goldenberg. (2025). *HDP-Flow: Generalizable Bayesian Nonparametric Model for Time Series State Discovery*. The 41st Conference on Uncertainty in Artificial Intelligence. (UAI 2025).

[C11] Y. Turura , S. L. Friedman, A. Cremer, M. Maddah, **S. Tonekaboni**. (2025).
The Latentverse: An Open-Source Benchmarking Toolkit for Evaluating Latent Representations. In Conference on Health, Inference, and learning (CHIL 2025).

[J3] M. D. McCradden, K. Thai, A. Assadi, **S. Tonekaboni**, I. Stedman, S. Joshi, M. Zhang, F. Chevalier, A. Goldenberg. (2025). *What makes a good decision with artificial intelligence? A grounded theory study in paediatric care*. (BMJ Evidence-Based Medicine).

[C10] C. Wang, S. Gupta, X. Zhang, **S. Tonekaboni**, S. Jegelka, T. Jaakkola, C. Uhler. (2025). *An Information Criterion for Controlled Disentanglement of Multimodal Data*. International Conference on Learning Representations (ICLR 2025).

[C9] S. Nagaraj, W. Gerych, **S. Tonekaboni**, A. Goldenberg, B. Ustun, T. Hartvigsen. (2025). *Learning under Temporal Label Noise*. International Conference on Learning Representations (ICLR 2025).

[J2] A. Nazaret*, **S. Tonekaboni***, G. Darnell, S. Ren, G. Sapiro, A. Miller (2023), *Modeling Personalized Heart Rate Response to Exercise and Environmental Factors with Wearables Data*. npj Digital Medicine. 6, 207

[C8] J. Yu, T. Behrouzi, K. Garg, A. Goldenberg, **S. Tonekaboni** (2023). *Dynamic Interpretable Change Point Detection for Physiological Data Analysis*. In Machine Learning for Health (ML4H 2023). (pp. 636-649). PMLR.

[C7] **S. Tonekaboni**, C.L Li, S. Arik, A. Goldenberg, T. Pfister. (2022). *Decoupling Local and Global Representations of Time Series*. In conference on Artificial Intelligence and Statistics (AISTAS 2022).

[C6] **S. Tonekaboni**, G. Morgenshtern, A. Assadi, A. Pokhrel, X. Huang, A. Jayarajan, R. Greer, G. Pekhimenko, M. McCradden, F. Chevalier, M. Mazwi, A. Goldenberg. (2022). *How to validate Machine Learning Models Prior to Deployment: Silent trial protocol for evaluation of real-time models at the ICU*. In Conference on Health, Inference, and learning (CHIL 2022).

[C5] A. Weatherhead, R. Greer, M. Moga, M. Mazwi, D. Eytan, A. Goldenberg, **S. Tonekaboni**. (2022). *Learning Unsupervised Representations for ICU Timeseries*. In Conference on Health, Inference, and learning (CHIL 2022).

[C4] **S. Tonekaboni**, D. Eytan, A. Goldenberg (2021). *Unsupervised Representation Learning for Time Series with Temporal Neighborhood Coding*. International Conference on Learning Representations (ICLR 2021).

[C3] **S. Tonekaboni**, S. Joshi, K. Campbell, D. Duvenaud, A. Goldenberg (2020). *What went wrong and when? Instance-wise feature importance for time-series black-box models*. Advances in Neural Information Processing Systems, 33. presented at the Conference on Neural Information Processing Systems (NeurIPS 2020).

[C2] **S. Tonekaboni**, S. Joshi, M.D. McCradden, A. Goldenberg (2019). *What Clinicians Want: Contextualizing Explainable Machine Learning for Clinical End Use*. *vol. 106 of Proceedings of Machine Learning Research* (PMLR). pp. 359-380.
presented at the Conference on Machine Learning for Healthcare (MLHC 2019).

[C1] **S. Tonekaboni**, M. Mazwi, P. Laussen, D. Eytan, R. Greer, S. Goodfellow, M. Brudno, A. Goldenberg. (2018). *Prediction of Cardiac Arrest from Physiological Signals in the Pediatric ICU*. *vol. 85 of Proceedings of Machine Learning Research* (PMLR). pp. 534-550.
presented at the Conference on Machine Learning for Healthcare (MLHC 2018).

[J1] H. Kassiri, **S. Tonekaboni**, N. Soltani, M. Tariqus Salam, J. Perez Velazquez, R. Genov. (2016). *Closed-Loop Neurostimulators: A Survey and A Seizure Predicting Design Example for Intractable Epilepsy Treatment*. IEEE Transactions on Biomedical Circuits and Systems

Peer Reviewed workshops/short papers:

C. Wang, S. Gupta, X. Zhang, S. Tonekaboni, S. Jegelka, T. Jaakkola, C. Uhler (2024), *An Information Criterion for Controlled Disentanglement of Multimodal Data*, Workshop on Unifying Representations in Neural Models NeurIPS 2024 . (Hounorable mention)

S. Nagaraj, W. Gerych, S. Tonekaboni, A. Goldenberg, B. Ustun, T. Hartvigsen (2024), *Time Series under Temporal Label Noise*, Workshop on Time Series in the Age of Large Models at NeurIPS 2024. (Spotlight)

G. Morgenshtern, A. Verma, S. Tonekaboni, R. Greer, J. Bernard, M. Mazwi, A. Goldenberg, and F. Chevalier (2023), *RiskFix: Supporting Expert Validation of Predictive Timeseries Models in High-Intensity Settings*, at EuroViz 2023

A. Nazaret, S. Tonekaboni, G. Darnell, S. Ren, G. Sapiro, A. Miller (2022), *Modeling Heart Rate Response to Exercise with Wearables Data*, In Learning from Time series in Health Workshop at NeurIPS 2022.

B. Naida, A. Weatherhead, S. Tonekaboni, A. Goldenberg (2022), *Continual Learning on Auxiliary tasks via Replayed Experiences: CLARE*, In Learning from Time series in Health Workshop at NeurIPS 2022.

S. Tonekaboni, S. Joshi, A. Goldenberg (2019), *I*ndividualized Feature Importance for Time Series Risk Prediction Models, In Machine Learning for Health Workshop at NeurIPS 2019.

M. McCradden, S. Tonekaboni, S. Joshi, A. Goldenberg (2019), *F*ive Pillars of Explainable Clinical Machine Learning, In Frontier of AI-Assisted Care (FAC) Scientific Symposium.

RESEARCH FUNDING

Utilizing high resolution physiological data and artificial intelligence to develop a pediatric cardiac arrest prediction tool for integration into bedside clinical practice
Agency: CIHR/NSERC - Role: Contributor - Amount: 950k (CAD) *2019*

PATENTS

Physiological predictions using machine learning [Link]
US 20240079112 A1 | Published: March 7, 2024
Inventors: A. Nazaret, S. Tonekaboni, G. Darnell, S. Ren, G. Sapiro, A. Miller

INVITED TALKS

Invited speaker - INFORMS Annual Meeting (Atlanta, USA) *October 2025*
Title: Leveraging multimodal data in healthcare applications.

Invited speaker - AI for Science Summit (Cambridge, UK) *November 2024*
Title: Decoupled Multimodal Representation Learning for clinical data

Guest lecturer - University of Toronto *October 2023*
Course: Topics in Machine Learning for Healthcare, Department of computer science
Title: Learning Representations for Time Series in Healthcare

Guest lecturer - University of Toronto *April 2022*
Course: Intelligent medicine and machine learning, Dalla Lana school of public health
Title: Deployment of Machine Learning tools in clinical practice

Invited speaker - Toronto Deep Learning Series (Toronto, Canada) *October 2018*
Title: Prediction of cardiac arrest from physiological signals

Invited speaker - Toronto Rehabilitation Institute research round (Toronto, Canada) *November 2018*
Title: Prediction models for longitudinal data

Invited lecturer - Vector Institute, endless summer school (Toronto, Canada) *November 2018*
Title: Time-series analysis and prediction

ACADEMIC SERVICE

Organizing committee:

- Foundation models for brain and body workshop - NeurIPS 2025
- Machine Learning for Healthcare symposium 2024
- Human-Centric Representation Learning workshop - AAAI 2024
- Learning from time series in health - NeurIPS 2023 (Founding co-chair), ICLR 2024, NeurIPS 2025

Scholarship selection committee at the Vector institute, Canada *March 2024*

Reviewer: Nature, NeurIPS (top reviewer recognition), ICLR, AISTATS, CHIL, MLHC, TMLR, ICML.

Area chair: CHIL conference, ML4H symposium

STUDENT SUPERVISION AND MENTORING

Student supervision during Postdoc at the Broad institute of MIT and Harvard
Yoanna Turura (published [C11]), Hector Salma, Majd Alafrange - Undergraduate students
Quixuan Jin, Chenyu Wang (published [C10]), Yuxin Xiao (under review [C15]) - PhD students

Student supervision during PhD at the University of Toronto
Addison Weatherhead (published [C5]), Xi Huang, Aslesha Pokhrel (published [C6]) - Undergraduate student
Kopal Garg, Jenny Yu, Tina Behrouzi (published [C8]) - Master students

Mentor at HER code camp
A computer science camp for senior high school students from underrepresented groups