Shahriar Noroozizadeh

snoroozi@cs.cmu.edu • (412)773-0240 • snoroozi.github.io • LinkedIn.com/in/snoroozi/

EDUCATION

Carnegie Mellon University, School of Computer Science and Heinz College *Pittsburgh, PA*
Doctor of Philosophy in Machine Learning and Public Policy **Expected: Dec 2026**
Advisors: Profs. George Chen, Jeremy Weiss (NIH), Zachary Lipton (ML)
GPA: 4.1/4.3

Master of Science in Machine Learning (Computer Science) **May 2022**
GPA: 4.1/4.3

Master of Science in Biomedical Engineering (Computational Neuroscience) **Dec 2020**
Advisor: Prof. Bin He
GPA: 4.0/4.0

University of British Columbia *Vancouver, Canada*
Bachelor of Applied Science – Engineering Physics (EE+CS Specialization) **May 2018**
Advisor: Prof. Septimiu Salcudean

SKILLS

Programming Languages: Python, C/C++, Java, MATLAB, Git, Bash
Deep Learning Libraries: PyTorch, TensorFlow, Transformers, HugginFace, OpenAI Gym
Python Libraries: Numpy, Matplotlib, Pandas, SciPy, Scikit-Learn
Machine Learning and Deep Learning Techniques: Representation Learning, Contrastive Learning, Temporal ML, Multimodal ML, Regression, Survival Analysis, Reinforcement Learning, BERT, GPT, RNN, LSTM, CNN, VAE, GAN

PHD RESEARCH

Carnegie Mellon University, Machine Learning Department and Heinz College *Pittsburgh, PA*
Representation Learning for Unstructured Time Series/Longitudinal Data **Sep 2021 – Present**

Temporal-Supervised Contrastive Learning:
- Advanced supervised contrastive learning into the temporal domain through introducing a self-supervised framework learning hyperspherical embeddings centered on prototypical examples for enhanced intrinsic interpretability.
- Achieved state-of-the-art performance in clinical time series outcome prediction by underpinning predictive clustering.

Introducing an Alternative to Data Augmentation:
- Proposed a nearest neighbor pairing mechanism and validated it as an effective alternative for data augmentation in time series data to ensure realistic and meaningful contrastive learning when standard augmentations are not feasible.

Machine Learning Methodologies for Longitudinal Data:
- Devised novel representation learning techniques to adapt deep learning models for complexities of irregularly sampled and variable-length temporal data, crucial for deploying ML in high-stakes settings like clinical decision support in healthcare.

Multimodal Time Series Modeling:
- Investigating multimodal time series representation learning for grounding structured temporal data with textual information from clinical notes to boost the predictive performance of machine learning models in healthcare applications.

Generative Models for Event Trajectory Forecasting:
- Exploring the development of interpretable generative models to predict future event trajectories from historical data, enabling personalized event sequence analysis and time interval estimation.

SELECTED PROFESSIONAL EXPERIENCE

Sanofi *Cambridge, MA*
AI Research Scientist Intern, Unstructured Data Search Team **May – Aug 2024**

Microsoft *Vancouver, Canada*
Software Engineering Intern **Jan– May 2015**

SELECTED PUBLICATIONS

Shahriar Noroozizadeh, Jeremy C. Weiss, George H. Chen, Dec 2023, Temporal Supervised Contrastive Learning for Modeling Patient Risk Progression. In *Machine Learning for Health* (pp. 403–427). PMLR.

Alexander D. Kyriazis*, **Shahriar Noroozizadeh*,** Amir Refaee*, Woongcheol Choi*, Lap-Tak Chu*, [et al, including Septimiu E. Salcudean], 2019. An end-to-end system for automatic characterization of iba1 immunopositive microglia in whole slide imaging. *Neuroinformatics*, *17*, pp.373-389.

WORKSHOP PRESENTATIONS

Spandan Das*, Vinay Samuel***, Shahriar Noroozizadeh***, "TLDR at SemEval-2024 Task 2: T5-generated clinical-Language summaries for DeBERTa Report Analysis", NAACL – Proceedings of the 18th International Workshop on Semantic Evaluation, Mexico City, Mexico, June 2024. [Poster]

Shahriar Noroozizadeh, Jeremy C. Weiss, George H. Chen, "Temporal Supervised Contrastive Learning with Applications to Tabular Time Series Data", AAAI – Representation Learning for Responsible Human-Centric AI workshop, Washington DC, USA, 13 Feb 2023. [Oral Presentation]

Ye Won Byun*, Cathy Jiao***, Shahriar Noroozizadeh***, Jimin Sun*, Rosa Vitiello*, "ET tu, CLIP? Addressing Common Object Errors for Unseen Environments", CVPR – Embodied AI workshop, New Orleans, USA, 19 June 2022. [Poster]

ADDITIONAL SELECTED RESEARCH EXPERIENCE

University of British Columbia *Vancouver, Canada*

Student Researcher, Robotics and Control Laboratory, Prof. Septimiu Salcudean **Sep 2016 – Aug 2018**

- **Machine Learning in Pathology Image Analysis:** Collaborated in a team of three to developed machine learning software for automatic segmentation and analyzing brain images post-TBI, using techniques like Random Forest and CNNs.
- **Real-Time Breast Cancer Screening Software:** Developed a GPU-accelerated algorithm for photoacoustic tomography, employing CUDA for GPU-accelerated real-time 3D visualization, significantly reducing diagnosis time.

University of Southern California *Los Angeles, CA*

Research Assistant, ODDS Research Group, Prof. Meisam Razaviyayn **May 2017 – Aug 2017**

- **RNA Sequence Clustering Optimization:** Devised unsupervised machine learning algorithms for pre-clustering Pacific Biosciences RNA sequences, boosting accuracy and efficiency in the CONVEX tool for transcriptome error correction.

SELECTED ACADEMIC PROJECTS

Clinical Trial NLI Enhancement – SemEval 2024 Task 2 – Advanced NLP (CMU) **Nov 2023 – Mar 2024**

- Designing a model for SemEval 2024's Safe Biomedical NLI task, focusing on enhancing LLMs for reliable inference in clinical trials' data analysis.
- Utilizing GPT for generating coherent premises and finetuning DeBERTa on entailment and contradiction classification.
- Implemented noisy channel model prompting enhancing few-shot learning performance for sparse and unbalanced data.

"Jarvis, Follow my Lead" – Multi-Modal Machine Learning (CMU) **Jan – Jun 2022**

- Spearheaded a team project to align natural language instructions with egocentric visual cues for household task execution using ALFRED benchmarks.
- Improved model robustness by incorporating pre-trained CLIP encoders with Episodic Transformer architecture.
- Showcased the project's advancements at the CVPR Embodied AI Workshop, highlighting improvements in object recognition and directive interpretation.

"BERT, do you still love me?" – Probabilistic Graphical Models (CMU) **Jan – May 2021**

- Conducted a comprehensive study on enhancing BERT and RNNs with graphical models for POS tagging and NER tasks, demonstrating improved encoding capabilities.
- Demonstrated a performance boost by integrating a CRF layer with a pre-trained BERT through end-to-end training.
- Awarded Best Project and Poster in the PGM course for innovative approach to NLP task performance enhancement.

Model-Based RL with PETS – Deep Reinforcement Learning (CMU) **Nov – Dec 2019**

- Implemented the Probabilistic Ensemble and Trajectory Sampling (PETS) algorithm in the OpenAI-Gym Pusher environment, enabling a robotic arm to push objects to target locations using a probabilistic neural network ensemble.
- Employed model predictive control with the cross-entropy method for effective planning and trajectory sampling, enhancing the algorithm's decision-making capabilities.

AWARDS & HONOURS

Tata Consultancy Services (TCS) Presidential Fellowship	**Sep 2024**
Natural Sciences and Engineering Research Council of Canada (NSERC) 3-year Doctoral Fellowship	**Sep 2023**
Suresh Konda Memorial Best First Research Paper Award	**May 2023**
Center for Machine Learning and Health Digital Health Innovation Fellowship	**Sep 2022**
Carnegie Mellon University Presidential Fellowship	**Sep 2018**
The Google Games at UBC: **3rd** Place / **1st** Place in Coding Challenge, Google	**May 2016**
IEEEXtreme 9.0 Programming Contest: **1st** Place at UBC, **4th** in Canada, IEEE	**Oct 2015**

PROFESSIONAL SERVICES AND TEACHING EXPERIENCE

Reviewer: ICLR 2024,2025, CHIL 2024, NeurIPS 2023, AAAI 2023

Teaching Assistant: Probabilistic Graphical Models (2024), Machine Learning for Problem Solving (2023), Unstructured Data Analytics (2023,2022), PhD Microeconomics(2024–2022), Neural Signal Processing (2020), Algorithms and Data Structures (2017, 2018)