

Monireh Safari

Highlighted Skills

ML researcher with 7+ years in deep learning (PyTorch, TensorFlow), specializing in transformer-based foundation models, LLMs, and self-supervised learning on biological data. Proven track record building scalable pipelines and translating research into healthcare applications.

Education

- Sept 2022 - **University of Waterloo, Ontario, Canada**
Sept 2026 PhD Student in Computer Science – **GPA: 95/100**
Supervisor: Prof. Lila Kari
- Sept 2022 **Sharif University of Technology, Tehran, Iran**
M.Sc. in Computer Engineering - **GPA: 99/100**
- Sept 2019 **Amirkabir University of Technology, Tehran, Iran**
B.Sc. in Computer Engineering - **GPA: 93/100**

Awards

- | | | |
|------|--|---|
| 2026 | Exceptional Flash Talk Award | <i>Canadian Bioinformatics Hub Conference</i> |
| 2026 | WiCS Graduate Student Scholarship - CA\$3K | <i>University of Waterloo</i> |
| 2026 | David R. Cheriton Graduate Scholarship - CA\$20K | <i>University of Waterloo</i> |
| 2025 | AMII Travel Bursary for AI Conference - CA\$1.5K | <i>University of Waterloo</i> |
| 2024 | Cotton Family Women in Mathematics Scholarship - CA\$3.5K | <i>University of Waterloo</i> |
| 2022 | Doctoral Entrance Award for Women in Math - CA\$5K | <i>University of Waterloo</i> |

Experience

- Sept 2022 – **Kari Genomic Lab – University of Waterloo, Waterloo, Ontario, Canada**
Present **Graduate Research Assistant**
- Developed transformer-based DNA foundation models (BERT-style) with self-supervised learning on large-scale genomic datasets (1M+ sequences), demonstrating LLM-style pretraining and fine-tuning workflows applicable to biological sequence data; achieved 15% improvement in taxonomic classification.
 - Designed *BarcodeMAE*, a scalable masked autoencoder foundation model for biological sequences; demonstrated 10.5% performance improvement and strong generalization across multiple taxonomic cohorts.
 - Uncovered environmental genomic signatures across hundreds of microbial genomes using alignment-free ML, generating novel scientific hypotheses on how extreme conditions override taxonomic patterns.
- Jan 2025 – **Princess Margaret Cancer Research Center - Vector Institute, Toronto, Ontario, Canada**
April 2025 **Research Intern**
- Designed and trained a probabilistic Adversarial Variational Autoencoder in PyTorch to learn interpretable representations from single-cell clinical methylation data, informing downstream biomarker discovery.
- Sept 2020 – **Sharif University of Technology, Tehran, Iran**
May 2022 **Graduate Research Assistant**
- Developed a graph neural network model for cancer-related gene identification by integrating biological network structure with genomic features, achieving 4% improvement.
- Jun 2018 – **Tebinja Company, Tehran, Iran**
Dec 2018 **Machine Learning Developer**
- Designed and deployed an interpretable disease prediction system using Bayesian Networks and Deep Q-Networks in Python, achieving 86% accuracy on clinical symptom data.

Skills

Languages: Python, JAVA, C/C++, MATLAB/Octave, R, Go, SQL
Frameworks: PyTorch, TensorFlow, OpenCV, React, Node.js, JavaScript, Django, Flask
Tools: Git, L^AT_EX, MEGA, MySQL, MongoDB

Publications and Preprints

- 2026 **Safari, M.**, et al. (2026). BarcodeBERT: Transformers for biodiversity analyses. **Bioinformatics Advances**, vbag054. <https://doi.org/10.1093/bioadv/vbag054>
- 2025 **Safari, M.**, Butler, J., Randhawa, G. S., Hill, K. A., & Kari, L. (2025). Life at the extremes: Maximally divergent microbes with similar genomic signatures linked to extreme environments. **Nucleic Acids Research: Genomics and Bioinformatics**. <https://doi.org/10.1093/nargab/lqaf189>
- 2025 **Safari, M.**, Arias, P. M., Lowe, S. C., Kari, L., Chang, A. X., & Taylor, G. W. (2025). Enhancing DNA Foundation Models to Address Masking Inefficiencies. [arXiv:2502.18405](https://arxiv.org/abs/2502.18405). <https://arxiv.org/abs/2502.18405>

Posters and Presentations

- 2026 Machine Learning Reveals an Extreme-Environment k-mer-based Watermark Imprinted Across Eukaryotic Genomes. **Safari, M.**, et al - **Best Flash Talk Award - Canadian Bioinformatics Hub Conference 2026**
- 2025 Enhancing DNA Foundation Models to Address Masking Inefficiencies. **Safari, M.**, et al - AI4NA Workshop and MLGenX Workshop - **ICLR 2025**
- 2025 Adversarial Learning for Scalable High-Resolution Cancer Lineage Tracing from Single-Cell Methylation Data. **Safari, M.**, et al - **Vector Remarkable Research Symposium 2025**
- 2025 Evidence for a Striking Convergence of Extreme Environment-Associated Genomic Signatures. Butler, JW., **Safari, M.**, et al - **EMGS Conference 2025**
- 2024 From Extreme Life to Genomic Insights: Optimized parameters for extremophile genomic signature classification. **Safari, M.**, et al - **EMGS Conference 2024**
- 2024 Optimized machine learning methods permit the discovery of an environmental component of genomic signatures. Butler, JW., **Safari, M.**, et al - **EMGS Conference 2024**
- 2023 BarcodeBERT: Transformers for Biodiversity Analysis. Millán Arias, P., Sadjadi, N., **Safari, M.**, et al - Workshop on Self-Supervised Learning - **NeurIPS 2023**

Talks and Media

- Feb 2026 **The Rising Star Seminar Series**
Invited Speaker
Presented PhD research on DNA foundation models and computational genomics to undergraduate students, discussing research pathways and career opportunities in computational biology.
- Jan 2026 **The Western Biology & Evolution Seminar Series (WBESS)**
Invited Podcast Guest
Discussion of the paper: Life at the extremes: maximally divergent microbes with similar signatures

Honors

- | | | |
|------|--|---|
| 2022 | Top 10 in Waterloo AI Fall 2022 Data Challenge | <i>University of Waterloo</i> |
| 2021 | Ranked 1st among M.Sc. Students in Computer Engineering | <i>Sharif University of Technology</i> |
| 2019 | Qualified to direct admission to M.Sc. program in ECE | <i>Sharif University of Technology</i> |
| 2019 | Qualified to direct admission to M.Sc. program in ECE | <i>Amirkabir University of Technology</i> |
| 2019 | Ranked 3rd among B.Sc. Students in Computer Engineering | <i>Amirkabir University of Technology</i> |
| 2013 | Best Technical Presentation in RoboCup 2013 | <i>Eindhoven, Netherlands</i> |

Teaching Experience

- Sept 2022 - **University of Waterloo**, Ontario, Canada
Present *Invited Lecturer, Teaching Assistant and Instructional Apprentice, Computer Science Department*
Course: Logic and Computation