Samaneh H. Moghaddam

437 808 2111, samaneh.moghaddam@utoronto.ca

SUMMARY

Innovative AI researcher and postdoctoral fellow with 10+ years of experience in artificial intelligence, machine learning, and multimodal data analysis, combining computational social science, NLP, and applied modeling. I have designed and implemented end-to-end supervised learning pipelines, including multimodal fusion models that integrate textual, behavioral, and relational signals for tasks such as abusive language and social bot detection. I have extensive experience in large-scale data collection, curation, preprocessing, and quality assurance, as well as supervising research assistants and leading cross-disciplinary projects. My work emphasizes robustness, generalizability, scalability, and operational relevance, with outcomes published in peer-reviewed journals (IEEE TDSC, Computers in Human Behavior Reports) and conferences and includes publicly available facets such as AbuseBERT. I have also led industrial R&D projects deploying predictive systems for credit scoring, demonstrating practical implementation skills.

EDUCATION

Postdoc | Faculty of Information, University of Toronto, Toronto, Canada — Present

- Title of project: Abusive language detection for identifying cyber-violence against professionals serving the public
- Supervisors: Dr. Kelly Lyons, Dr. Cheryl Regehr, and Dr. Frank Rudzicz

Ph.D. | Computer Science and Engineering Department, Shahid Beheshti University, Tehran, Iran — 2022

- Dissertation: Detecting fake accounts controlled by social bots
- Supervisor: Dr. Maghsoud Abbaspour
- Awarded **Iran National Elite Foundation Research Scholarship**
- Achieved **A+ grade** in dissertation defense

M.Sc. | Department of Computer Engineering – Tarbiat Modares University, Tehran, Iran — 2007

- Thesis: Software watermarking for software license verification
- Supervisor: Dr. Saeed Jalili
- Awarded **Telecommunication Research Center (ITRC) Research Grant**

B.Sc. | Department of Computer Engineering - Ferdowsi University, Mashhad, Iran — 2004

- Final Year Project (Dissertation): Processing Persian sentences to extract keywords in search engines
- Supervisor: Dr. Mohsen Kahani
- Honored as **top outstanding students**

SELECTED PEER REVIEWED PUBLICATIONS

- **H. Moghaddam, S.,** Regehr, C., Lyons, K., Goel, V., & Regehr, K. (2025). Drivers of online abuse against Canadian public health officials: An LLM-based temporal analysis. **Canadian Journal of Public Health**. (Under Review)

- Regehr, C., Regehr, K., **H. Moghaddam, S.,** Lyons, K., Goel, V. (2025). Workplace violence in an online era: Technology-facilitated threats and harassment of public health. **European Congress of Psychiatry**. (Under Review)
- **H. Moghaddam, S.,** Lyons, K., Regehr, C., Goel, V., & Regehr, K. (2025). Towards a comprehensive taxonomy of online abusive language informed by machine learning. **Computers in Human Behavior Reports.**
- **H. Moghaddam, S.,** Lyons, K., Rudzicz, F., Regehr, C., Goel, V., & Regehr, K. (2025). Enhancing machine learning in abusive language detection with dataset aggregation. In **Proceedings of the 35th IEEE International Conference on Collaborative Advances in Software Computing** (CASCON).
- **H. Moghaddam, S.,** & Abbaspour, M. (2022). Friendship preference: Scalable and robust category of features for social bot detection. **IEEE Transactions on Dependable and Secure Computing**, 20(2), 1516–1528.
- **H. Moghaddam, S.,** & Abbaspour, M. (2020). Economic feasibility study for stealthy socialbot network establishment**. IET Information Security**, 14(6), 694–703.
- Khademi, M., **H. Moghaddam, S.,** & Abbaspour, M. (2017). An empirical study of the effect of profile and behavioral characteristics on the infiltration rate of socialbots. **In Proceedings of the Iranian Conference on Electrical Engineering (ICEE)**. IEEE.
- **H. Moghaddam, S.,** & Abbaspour, M. (2014). Sensitivity analysis of static features for Android malware detection. **In Proceedings of the 22nd Iranian Conference on Electrical Engineering (ICEE).** IEEE.

RESEARCH EXPERIENCE

Abusive language detection and driver of cyber-violence 2023-Present

- Developed a hierarchical taxonomy of abusive language to integrate multi-labeled datasets, bias mitigation, and enhance generalizability of models.
- Applied progressive learning to adapt models to novel linguistic behaviors
- Collected and analyzed 2M+ tweets targeting public health professionals, investigating correlations between abusive trends and policy or environmental factors.
- Emphasized interdisciplinary approaches by linking computational methods with social phenomena, leading to temporal analyses of the drivers of cyberviolence.
- Developing multi-modal graph-aware language models that integrate textual, behavioral, and relational signals within heterogeneous social networks for robust detection of subtle abuse.
- Publication in reputable interdisciplinary venues such as *Computers in Human Behavior Reports.*

Detecting fake accounts controlled by social bots, 2016-2022

- Developed and curated 256 AI-driven agents (social bots) each with a unique infiltration strategy to gain human trust and social influence online.
- Proposed a novel digital footprint representation, leading to robust, scalable and generalizable features for trustworthy bot detection.
- Developed standalone and multi-expert classification models by decision fusion.
- Proposed an analytical framework for understanding the economic feasibility of stealthy bot networks.
- Published in *IEEE Transactions on Dependable and Secure Computing (TDSC),* a top-tier journal in security and trustworthy computing.

Earlier Work in Network and Behavioral Analysis 2012-2015

- Performed anomaly detection in mobile networks using graph mining and semi-supervised ML, integrating social network analysis concepts.

- Conducted sensitivity analysis for Android malware detection and developed software watermarking algorithms, combining technical rigor with practical insights.

Research Interest

- **Multimodal Machine Learning and Representation Fusion**: Exploring how diverse digital footprints, textual, behavioral, and relational, can be recombined or fused to improve classification, and robustness of complex data, with growing interest in extending these methods to medical imaging and clinical text.
- **Applied Machine Learning**: Developing fair, generalizable, and scalable models that bridge computational methods and domain knowledge in health and social data contexts.
- **Context-Grounded Modeling:** Exploring how contextual cues anchor learning to real-world patterns, enabling models to make more interpretable and contextually informed predictions across domains.

SKILLS and Theories

- **AI & NLP:** Machine Learning, Deep Learning, Natural Language Processing, Large Language Models, Adaptive and Fair Learning
- **Programming & Tools:** Python, PyTorch, PyG, Transformers, Scikit-learn, Data Analysis & Visualization
- **High-Performance & Cluster Computing:** Familiar with UofT HPC environment, including HCPI cluster, and Slurm job scheduling.
- **Research & Leadership:** Project leadership and cross-functional collaboration, mentoring researchers, and translating real-world, operational requirements into scalable, deployable AI solutions.

PAPER REVIEW & PROGRAM COMMITTEE SERVICE

- **Reviewer,** Computational Intelligence — 2025
- **Program Committee Member**, CASCON Artifacts Track — 2025
- **Reviewer**, Computational Intelligence Journal — 2024
- **Reviewer**, 11th Conference on Information and Knowledge Technology — 2020
- **Reviewer**, 4th International Conference on Applied Research in Computer and IT — 2016
- **Reviewer**, 19th National CSI Computer Conference — 2014

Mentoring and Supervision

- **Technical Facilitator, Agentic AI Bootcamp, Vector Institute** — 2025-2026
  - Delivered workshops on LLM-based agent frameworks and supervised hands-on coding projects.
- Mentored participants on AI experimentation, prompt engineering, and evaluation of agentic models
- **Co-supervised Research Assistants, University of Toronto**— 2024-2025
  - Co-supervised two RAs on abusive language detection project, focusing on large-scale data collection, preprocessing, and dataset integration.
- **Co-supervised M.Sc. Students, Shahid Beheshti University** — 2016-2022

- M.Sc. on projects including:
  - Managing Trust in Social Internet-of-Things Systems
  - Anomaly detection in online social networks
  - Security and Intrusion Detection in 4G Mobile Networks
  - Behavioral Modeling for Mobile Botnet Detection

LECTURESHIP

- **Software Engineering,** Computer Science Department, Shahid Beheshti University, Jan. 2015
- **Computation Theory,** Computer Science Department, Shahid Beheshti University, Jan. 2013
- **Programming Foundation,** Computer Science Department, Shahid Beheshti University, Jan. 2012

HONORS AND AWARDS

- Recognized as a scientific elite by the **Iran National Elite Foundation**, 2015.
- Achieved the membership of **Brilliant talents of Shahid Beheshti University**, 2013.
- First place in fraud detection league in **Amir Kabir International Robotics and Artificial Intelligence Contests**, 2012
- First place in artificial intelligence competition in **Amir Kabir International Robotics and Artificial Intelligence Contests**, 2012.
- **Iranian Society of Cryptology** membership, 2007.
- Chosen in the second stage of **National Computer Olympiad,** 1999.
- Second place in student section of **Kharazmi festival** for "A Euclidean geometry model for non-Euclidean geometry", 1998.
- Admitted in the entrance exam of **National Organization for Development of Exceptional Talents** (NODET), 1996.