Farima Fatahi Bayat

Email | Website | LinkedIn | GitHub | Google Scholar

RESEARCH INTERESTS

My research focuses on **advancing responsible AI**, with a particular emphasis on **mitigating LLM factuality and reasoning hallucinations**. I have experience developing reference-based evaluation frameworks to identify factuality challenges and designing adaptive intervention and calibration techniques to improve LM truthfulness. My industry experience includes working with **knowledge graphs** for question answering, building **recommendation systems** for product ranking, and recently, evaluating failures of tool-augmented LLMs in mathematical reasoning.

EDUCATION

University of Michigan

Ph.D. in Computer Science & Engineering 2022 - 2025 (expected)
Advisor: H. V. Jagadish
Research Collaborator: Lu Wang

M.Sc. in Computer Science & Engineering 2020 - 2022
GPA: 3.9 / 4.0

University of Tehran

B.Sc. in Computer Engineering 2014 - 2018

PROFESSIONAL EXPERIENCE

University of Michigan

August 2020 - present
Graduate Student Research Assistant

- Advisors: H. V. Jagadish & Lu Wang
- Developed a factuality evaluation framework with a real-world benchmark and metric challenging LMs in factual generation. (ACL 2025 main)
- Enhanced LM truthfulness at inference via adaptive intervention and uncertainty. (ACL Findings 2024)
- Built activation-based calibration and confidence-guided decoding methods to improve factuality. (EMNLP 2024)
- Proposed an open information extraction pipeline that extracts compact facts from text. (NAACL 2022)

Megagon Labs. - Compound AI System team

Summer 2025
Research Scientist Intern
Mentor: Pouya Pezeshkpour

- Evaluating LM reasoning hallucinations caused by over-reliance on external tools. (**ongoing**)

Apple Inc. - Knowledge Platform team

January 2023 - September 2023
Machine Learning Research Intern
Manager: Yunyao Li

- Built a system to verify and revise LM outputs using external knowledge. (EMNLP 2023: System Demonstrations)
- Proposed a prompt engineering tool that refines prompt demonstrations through active learning. (APE, DaSH 2024)

Digikala Inc. - Data Science team

Summer 2019
Data Engineer Intern

- Developed a recommendation system that ranks related items to each product on the website.
- Built a user purchase prediction model for real-time session analysis on e-commerce platforms.

SELECTED PUBLICATIONS

◇ **F. Fatahi Bayat**, L. Zhang, S. Munir, and L. Wang, "FactBench: A Dynamic Benchmark for In-the-Wild Language Model Factuality Evaluation," *Association for Computational Linguistics,* 2025. ***Oral Presentation.***

◇ X. Liu, **F. Fatahi Bayat**, L. Wang, "Enhancing Language Model Factuality via Activation-Based Confidence Calibration and Guided Decoding," *Empirical Methods in Natural Language Processing,* 2024. [code]

◇ **F. Fatahi Bayat**, X. Liu, H. Jagadish, and L. Wang, "Enhanced Language Model Truthfulness with Learnable Intervention and Uncertainty Expression," *Association for Computational Linguistics (Findings),* 2024. [code]

◇ K. Qian, Y. Sang, **F. Fatahi Bayat**, A. Belyi, X. Chu, Y. Govind, S. Khorshidi, R. Khot, K. Luna, A. Nikfarjam, X. Qi, F. Wu, X. Zhang, and Y. Li, "APE: Active Learning-based Tooling for Finding Informative Few-shot Examples for LLM-based Entity Matching," *DaSH,* 2024.

◇ **F. Fatahi Bayat**, K. Qian, B. Han, Y. Sang, A. Belyi, S. Khorshidi, F. Wu, I. F. Ilyas, and Y. Li, "FLEEK: Factual Error Detection and Correction with Evidence Retrieved from External Knowledge," *Empirical Methods in Natural Language Processing (Demonstration Track),* 2023.

◇ **F. Fatahi Bayat**, N. Bhutani, and H. Jagadish, "CompactIE: Compact Facts in Open Information Extraction," *North American Chapter of the Association for Computational Linguistics,* 2022. [code]

◇ H. Mahdiani, A. Khadem, A. Ghanbari, M. Modarressi, **F. Fatahi Bayat**, and M. Daneshtalab, "ΔNN: Power-efficient Neural Network Acceleration using Differential Weights," *IEEE Micro,* 2019.

TEACHING EXPERIENCE

University of Michigan - Electrical Engineering and Computer Science Department
Graduate Student Instructor

- Delivered advanced NLP lectures on: **RAG**, **Prompting**, and **Information Extraction**
- "EECS 486: Information Retrieval and Web Search" by Rada Mihalcea — Winter 2025
- "EECS 595: Natural Language Processing" by Joyce Chai & David Jurgens — Fall 2023 - Fall 2024

University of Tehran - Electrical and Computer Engineering Department
Teaching Assistant

- "Computer Architecture" by Saeid Safari — Winter 2017
- "Operating Systems" by Mehdi Kargahi — Fall 2017
- "Computer Aided Digital System Design" by Mehdi Modarressi — Winter 2018
- "Computer Networks" by Ahmad Khonsari — Winter 2018

PROFESSIONAL SERVICES

Area Chair and Reviewer in ARR (ACL Rolling Review) — 2024
Reviewer in ICLR (International Conference on Learning Representations) — 2024
Reviewer in VLBD Journal (The International Journal on Very Large Data Bases) — 2023
Reviewer in Computer Speech and Language Journal — 2023

AWARDS AND HONORS

Rackham Conference Travel Grant. — 2022, 2024
Rackham Doctoral Intern Fellowship, University of Michigan. — 2023
Awarded M.Sc. admission as an exceptional talent student, University of Tehran. — 2018