Hamidreza Amirzadeh

CE M.Sc. STUDENT , SHARIF UNIVERSITY OF TECHNOLOGY, TEHRAN

✉ hamid.amirzadeh78@sharif.edu | 🏠 hamid-amir.github.io | 🖧 hamid-amir | 💼 HamidrezaAmirzadeh | 🎓 Hamidreza Amirzadeh

Education

Sharif University of Technology
Tehran, Iran

M.Sc. in Artificial Intelligence, Computer Engineering
Sept 2022 - Present

- GPA: 3.93/4
- Thesis Title: Transformers interpretability on natural language tasks
- Advisor: Hossein Sameti
- Courses: Machine Learning, Deep Learning, Large Language Models, Natural Language Processing, Security and Privacy in Machine Learning, Convex Optimization, Digital Signal Processing

Tehran Polytechnic (Amirkabir University of Technology)
Tehran, Iran

B.Sc. in Mechanical Engineering
Sept 2017 - Sept 2021

- GPA: 3.65/4
- Thesis Title: Gait phase detection using machine learning algorithms
- Advisor: Mohammad Zareinejad
- Courses: Numerical Calculations, Signals and Systems, Linear Control systems

National Organization for Development of Exceptional Talents
Kerman, Iran

Diploma in Mathematics and Physics Discipline
Sept 2014 - Jun 2017

- GPA: 19.34/20

Research Interests

- Mechanistic Interpretability
- Vision LMs
- Security and Privacy in AI
- Trustworthy AI
- Robust NLP
- Deep Learning

Publications

4. data2lang2vec: Data Driven Typological Features Completion [paper] [code]
COLING 2025 (Under Review)

Hamidreza Amirzadeh, Sadegh Jafari, Anika Harju, Rob van der Goot

3. How Language Models Prioritize Contextual Grammatical Cues? [paper] [code]
BlackboxNLP @ EMNLP 2024

Hamidreza Amirzadeh, Afra Alishahi, Hosein Mohebbi

2. HalluSafe at SemEval-2024 Task 6: An NLI-based Approach to Make LLMs Safer by Better Detecting Hallucinations and Overgeneration Mistakes [paper] [code]
SemEval 2024

Zahra Rahimi, **Hamidreza Amirzadeh**, Alireza Sohrabi, Zeinab Taghavi, Hossein Sameti

1. Imaginations of WALL-E : Reconstructing Experiences with an Imagination-Inspired Module for Advanced AI Systems [paper]
arXiv preprint 2023

Zeinab Taghavi, Soroush Gooran, Seyed Arshan Dalili, **Hamidreza Amirzadeh**, Mohammad Jalal Nematbakhsh, Hossein Sameti

Research Experience

Remote collaborator at Tilburg University
Tilburg, Netherlands

Under Supervision of Prof. Afra Alishahi, and Hosein Mohebbi, Ph.D
Apr - Sep 2024

- Investigated how different language models use multiple contextual cues to disambiguate grammatical tasks. (BlackboxNLP @ EMNLP 2024)

Remote collaborator at IT University of Copenhagen, NLPnorth research group
Copenhagen, Denmark

Under Supervision of Prof. Rob van der Goot
Feb - Sep 2024

- Developed a more complete lang2vec system to improve multilingual models performance. (COLING 2025)

Graduate Research Assistant at Sharif University of technology *Tehran, Iran*

Under Supervision of Prof. Hossein Sameti *Sept 2022 - Present*

- Working on interpretability of transformer based language models and specificly quantifying token attribution in Transformers.
- Proposed an imagination inspired module to enhance the performance of multi-modal LLMs on text only tasks.
- Particpated in a shared task of SemEval 2024 titled SHROOM, a Shared-task on Hallucinations and Related Observable Overgeneration Mistakes. (SemEval 2024)

NLP Research Intern at ASR GOOYESH PARDAZ *Tehran, Iran*

Under Supervision of Prof. Hossein Sameti *Summer 2023*

- Developed a novel Persian knowledge graph.

Undergraduate Research Assistant at Tehran Polytechnic *Tehran, Iran*

Under Supervision of Dr. Mohammad Zareinejad *Spring 2022*

- Gait event detection using inertial sensors and machine learning algorithms. [code]

Teaching Experience

Teaching Assistant *Tehran, Iran*

Sharif University of technology *Sept 2022 - Present*

- **Artificial Intelligence**, Prof. Mohammad H. Rohban, Spring 2023, Fall 2023, Spring 2024
- **Convex Optimization**, Prof. Amir Najafi, Spring 2024
- **Deep Learning**, Prof. Mahdieh Soleymani, Spring 2024, Prof. Hamid Beigy, Fall 2023
- **Machine Learning**, Prof. Mehdi Jafari, Fall 2024
- **Security and Privacy in Machine Learning**, Prof. Amir M. Sadeghzadeh, Spring 2024
- **Natural Language Processing**, Prof. Ehsaneddin Asgari, Spring 2024

Honors & Awards

2024	**Participating as a reviewer**, for SemEval 2024	*Iran*
2023	**Second place in Rahisho competition**, a problem-oriented event held by Iran's National Elites Foundation	*Iran*
2023	**Participating as a reviewer**, for International Conference of Mechatronics and Robotics (ICRoM) [certificate]	*Iran*
2022	**Ranked in top of 0.05%**, in the National Entrance Exam for M.Sc. in Computer Engineering	*Iran*
2017	**Ranked in top of 0.4%**, in the National Entrance Exam for B.Sc. of Iran among more than 148,000 students	*Iran*

Skills

Programming	Python (PyTorch, sklearn, NumPy, Pandas, Tensorflow), C/C++, MATLAB, R, SQL.
Tools and Frameworks	PyTorch, Keras, Tensorflow, HuggingFace, Transformers, Git, HTML/CSS, Latex.
Languages	Persian: Native, English: Fluent (TOEFL Score: **102/120**, Reading:28, Listening:27, Speaking:23, Writing:24)

Projects

Large Language Models course assignments @ Sharif University of Technology *Fall 2023*

PEFT methods, In-context Learning, VLMs, RLHF, DPO (PyTorch)

- Taught by Prof. Mahdieh Solyemani, Prof. Mohammad H. Rohban and Prof. Ehsaneddin Asgari

Security and Privacy in Machine Learning course assignments @ Sharif University of Technology *Spring 2023*

Adversarial attack methods, Defensive approaches, Data poisoning, Differential privacy (PyTorch)

- Taught by Prof. Amir M. Sadeghzadeh

Adversarial Persian QA *Fall 2022*

Robustness of Persian Transformer-based QA Models to Common Adversarial Attacks

- Part of our final project in the NLP course @ SUT

References

- **Afra Alishahi**, a.alishahi@tilburguniversity.edu, *Tilburg University*
- **Rob van der Goot**, robv@itu.dk, *IT University of Copenhagen*
- **Ehsaneddin Asgari**, asgari@berkeley.edu, *Qatar Computing Research Institute*
- **Hossein Sameti**, sameti@sharif.edu, *Sharif University of Technology*