Hamidreza Amirzadeh

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

hamid.amirzadeh78@sharif.edu | hamidrezaamirzadeh.github.io | HamidrezaAmirzadeh | HamidrezaAmirzadeh | Hamidreza Amirzadeh

Education

Sharif University of Technology — *Tehran, Iran*

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

- GPA: 18.31/20
- Thesis Title: Transformers interpretability on natural language tasks
- Advisor: Hossein Sameti
- Courses: Machine Learning, Large Language Models, Deep Learning, Convex Optimization, Security and Privacy in Machine Learning, Natural Language Processing, Digital Signal Processing

Tehran Polytechnic (Amirkabir University of Technology) — *Tehran, Iran*

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

- GPA: 16.37/20
- 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

- Trustworthy AI
- Natural Language Processing
- Models Interpretability
- Deep Learning

Publications

Imaginations of WALL-E : Reconstructing Experiences with an Imagination-Inspired Module for Advanced AI Systems [paper] — *arXiv*

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

- Seeking to submit in TACL 2024

Research Experience

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.

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

Under Supervision of Prof. Hossein Sameti — *Summer 2023*

- Development of a novel Persian knowledge graph.

Undergraduate Research Assistant at Tehran Polytechnic — *Tehran, Iran*

Under Supervision of Prof. Mohammad Zareinejad — *Spring 2022*

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

Honors & Awards

2024	**Participating as a reviewer**, in the peer review process 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**, in the peer review process 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*

Teaching Experience

Teaching Assistant *Tehran, Iran*

Sharif University of technology *Sept 2022 - Present*

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

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).

Projects

Adversarial Persian QA *Fall 2022*

Robustness Investigation of Persian Transformer-based Retrieval Question Answering models against some common adversarial attacks

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

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

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

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

Natural Language Processing course assignments @ Sharif University of Technology *Fall 2022*

Including Tokeinization, Statistical methods, Transformers, Text classification, Text generation (PyTorch)

- Taught by Prof. Ehsaneddin Asgari

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

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

- Taught by Prof. Amir M. Sadeghzadeh

Deep Learning course assignments @ Sharif University of Technology *Fall 2022*

Including CNN, RNN, Attention mechanism, Deep generative models, Reinforcement learning (PyTorch)

- Taught by Prof. Hamid Beigy

References

- **Hossein Sameti** , sameti@sharif.edu, Sharif University of Technology
- **Mohammad Zareinejad** , mzareaut@ac.ir, Tehran Polytechnic (Amirkabir University of Technology)