

Mohammad Mahdi Salmani-Zarchi

M.Sc. Student in Artificial Intelligence, University of Tehran

 WebpageZ  LinkedIn -  Google Scholar -  GitHub
 m.salmani78@ut.ac.ir / momasal78@gmail.com

Research Interests

Large Language Models, Human Preference Alignment, Trustworthy & Verifiable AI, Reinforcement Learning, Instruction Following, LLM Reasoning, Evaluation & Social Behavior.

Skills

AI & ML: Machine Learning, Deep Learning, Statistical Inference, Reinforcement Learning
NLP & LLMs: PEFT (LoRA, QLoRA), RLHF, RAG, Natural Language Inference, Agentic LLMs
Frameworks & Systems: PyTorch, TensorFlow, vLLM, Distributed Training, MLflow, Optuna
Programming & Tools: Python, C++, JavaScript; Git, $\text{\LaTeX}$, Docker

Education



M.Sc. in Computer Engineering - Artificial Intelligence

GPA: 17.95 / 20

University of Tehran, Tehran, Iran

Sep 2023 – Present

- **Thesis:** The Impact of Preference Alignment on Large Language Models Instruction Following



B.Sc. in Electrical Engineering

GPA: 15.28 / 20

Isfahan University of Technology, Isfahan, Iran

Sep 2018 – Dec 2022

- **Final Project:** Android app for diagnosing musculoskeletal abnormalities

Publications

* indicates first author

MDP-GRPO: Stabilized Group Relative Policy Optimization for Multi-Constraint Instruction Following* **ACL 2026 Main Conference. DOI**

- Identified and formalized three failure modes of GRPO under discrete, low-dispersion reward settings: low-variance amplification, mean-centering blindness, and zero-variance collapse.
- Proposed MDP-GRPO, a stabilized RL framework for multi-constraint instruction following with verifiable rewards, improving both training stability and reliable constraint satisfaction.

Evaluating Role-Prompted LLM Responses on Autism-Related Theory-of-Mind Tasks*

Accepted for publication in **Computers in Human Behavior** (Q1 Journal, IF: 12.2) **Preprint**

- Proposed a DSM-5-grounded evaluation framework for assessing theory-of-mind reasoning in LLMs under autism-spectrum-inspired cognitive constraints.
- Analyzed how prompting strategies and alignment mechanisms affect the separability and consistency of simulated cognitive profiles across multiple ToM task types and statistical properties.

Graph-Based Representations for Dynamic Curriculum Learning Under review at the **AAAI**.

- Proposed GraCL, a dynamic curriculum-learning framework that jointly updates a predictive model and a graph-based surrogate of curriculum quality using training feedback.

Work Experience

Research Intern, ODCL Project

ESSEC Business School, Cergy, France

Supervisors: *Prof. Emiliano Traversi*, *Dr. Pegah Alizadeh*, *Dr. Massinissa Hamidi* May 2026 – Present

- Developed a dynamic curriculum learning framework that adapts task ordering during training via learnable edge-preference parameters, formulated as a graph-based optimization problem.

Research Assistant, NLP Lab

University of Tehran, Tehran, Iran

Supervisors: *Prof. Hesham Faili*, *Dr. M.J. Dousti*

Dec. 2024 – Present

- Conducted research on human preference alignment and role-playing in LLMs, contributing to experimental validation and peer-reviewed manuscripts.

Natural Language Processing Engineer

MCI Lab, R&D Department

MCI, Tehran, Iran

March 2025 – May 2026

- Deployed resource-efficient LLM pipelines using supervised fine-tuning and reinforcement learning, including vLLM-based inference and serving. Using MLflow for experiment tracking.

Research Assistant, Cognitive Science Lab

Supervisor: [Dr. Abdol-Hossein Vahabie](#)

University of Tehran, Tehran, Iran

Spring & Summer 2025

- Investigated theory-of-mind and social reasoning capabilities of LLMs using cognitively inspired experimental designs grounded in psychological and DSM-5 frameworks.

Teaching Assistant

Natural Language Processing

Teacher: [Prof. Hesham Faili](#)

University of Tehran, Iran

[GitHub page](#) - Fall 2025

- Designed a course project and hands-on workshop on fine-tuning LLMs using LoRA/PEFT, covering data preparation, tokenizer analysis, training, and standardized evaluation.

Large Language Models

Teachers: [Dr. M.J. Dousti](#), [Dr. Yadollah Yaghoobzadeh](#)

University of Tehran, Iran

Spring 2025

- Developed assignments on in-context learning and human preference alignment, introducing students to prompt engineering and reinforcement learning from human feedback.

Deep Learning

Teacher: [Dr. Ahmad Kalhor](#)

University of Tehran, Iran

Fall 2024, Spring 2025

- Designed assignments on image captioning, focusing on CNN-RNN architectures and attention mechanisms.

Honors & Academic Service

- **National M.Sc. Entrance Examination:** Achieved national rank 29 in Computer Engineering.
- **MCI Scholarship:** Competitive scholarship supporting M.Sc. research in AI.
- **Session Chair:** ACL 2026 Conference, Virtual Session.

Projects

- **Text2SQL Chatbot:** LLM-based system for translating queries into SQL and visualizing results using agentic workflows. [🔗](#)
- **Natural Language Inference Model Optimization:** Optimized RoBERTa and LLaMA models using LoRA and QLoRA for memory-efficient inference. [🔗](#)
- **Assessing Data Privacy in LLMs:** Evaluated Llama-3.2 against privacy threats including Membership Inference Attack and Jailbreaking (PAIR). [🔗](#)
- **Tumor Detection:** Detect and classify MRI images of the brain for the diagnosis of different brain tumors using a convolutional neural network. [🔗](#)
- **Machine Translation:** Compare LSTM and Transformer models for machine translation tasks. [🔗](#)
- **Selected Additional Projects:**
 - Monte-Carlo Localization & SLAM on Vector Robot [🔗](#)
 - Sentiment Analysis with N-grams [🔗](#)