

# Amir Mohseni

Maastricht, The Netherlands | amir.mohseni.work@gmail.com | +31 06 8641 6012

linkedin.com/in/amirreza-mohseni | github.com/Amir-Mohseni | huggingface.co/AmirMohseni

## Education

---

**MSc Advanced Computer Science** Oct 2026 – Sep 2027  
**University of Oxford**, Oxford, United Kingdom

- Incoming Master's student, with an intended focus on **Artificial Intelligence**.

**Bachelor of Computer Science** Sep 2023 – Jun 2026  
**Maastricht University**, Maastricht, The Netherlands

- **GPA:** 8.9 / 10.0
- Selected for the **MARBLE Research Program (Honours Track)**.
- Thesis: investigated how users seek legal advice from LLMs, studied overconfidence issues in RL-trained LLMs, and developed low-cost legal-query routing classifiers achieving **91%+ accuracy** at an estimated **1000× lower cost** than repeated frontier-LLM classification.
- Created the **ICPC Club** at Maastricht University; organized weekly training sessions and contests, and qualified for **NWERC** in all eligible years.
- **Key Courses:** Machine Learning, High-Performance Computing, Cloud Computing, Recommender Systems, Databases.

**Bachelor of Computer Engineering** Sep 2021 – May 2023  
**Shahid Beheshti University**, Tehran, Iran

- **GPA:** 17.8 / 20.0
- Teaching Assistant for **Discrete Mathematics**, **Advanced Programming**, and **Data Structures and Algorithms**.
- Qualified and participated in the **ICPC World Finals**.
- Transferred to Maastricht University to continue the BSc in Computer Science.

## Publications

---

**CurveBench: A Benchmark for Exact Topological Reasoning over Nested Jordan Curves** 2026  
Amirreza Mohseni, Mona Mohammadi, Morteza Saghafian, Naser Talebizadeh Saradari. *arXiv preprint arXiv:2605.14068*. arXiv

## Experience

---

**Student Researcher**, Turing – Remote Apr 2025 – Present

- Leading a team of student researchers focused on **LLM modeling, reinforcement learning, and evaluation**.
- Fine-tuning in-house models using reinforcement learning to improve **tool-use**, code-generation, and reasoning capabilities.
- Trained **reward models** to predict human preferences in educational tasks using EduArena preference data, improving preference prediction accuracy from 62% to 75%.
- Presented the **EduArena** research platform at ICML 2025.

**Generative AI Researcher**, Scale AI – Remote Sep 2024 – Apr 2025

- Collaborated with ML researchers and PhD students to fine-tune and benchmark open-source models including **Gemma 2** and **Llama 3**.
- Improved reasoning performance on benchmarks such as **Math-500**, **MMMU**, and **GPQA Diamond** using SFT and DPO methods inspired by **STaR** and **V-STaR**.
- Delivered executive reports to enterprise clients, summarizing experimental results, performance evaluations, and research insights.

- Collected and cleaned large-scale social-media data from company Facebook pages using a licensed API.
- Trained and evaluated regression models to analyze post engagement and social trend patterns in response to news events.

## Projects

---

### Professional Projects

- **EduArena:** Core contributor to a large-scale research crowdsourcing platform for evaluating LLMs on educational tasks. Implemented automated pipelines for model routing, preference-data collection, and RLVR data generation using entropy and model-disagreement signals.
- **Reasoning Router:** Developed a router that predicts whether a query should be handled with *thinking* or *non-thinking* inference, reducing token usage by nearly **50%** while maintaining performance within **5%** of the best model's accuracy. [Blog / Models](#)

### Academic Projects

- **Legal Query Routing for User-LLM Interactions:** Built a classifier pipeline for analyzing legal-advice-seeking behavior in real-world chatbot conversations. Developed a legal-topic taxonomy, compared LLM- and encoder-based methods, and investigated overconfidence in RL-trained models for legal classification, proposing a simple mitigation strategy to reduce overconfidence.
- **CurveBench RLVR Fine-Tuning:** Developed RLVR-style fine-tuning experiments for vision-language models on exact topological reasoning tasks, using structured rewards based on rooted-tree recovery from images of nested Jordan curves.

## Achievements

---

- **Ranked 482<sup>nd</sup>** among ~120,000 mathematics participants (**Top 0.4%**) in the National University Entrance Exam, Iran (2021).
- **ICPC World Finalist** (45<sup>th</sup> Annual, 2022) — ICPC Profile.
- **Bronze Medalist**, ACM-ICPC Regional Contest (2023).
- Reached the **Top 120** in the National Olympiad of Informatics, Iran (2020).

## Skills

---

**Programming Languages:** Python, C/C++, Java

**Machine Learning & AI:** PyTorch, NumPy, TRL, vLLM, verl, Hugging Face Transformers

**Databases & Cloud:** SQL, AWS, GCP Vertex AI

**Other:** Competitive programming, technical writing, research communication, team coordination