

Keivan Rezaei

Curriculum Vitae



8125 Paint Branch Dr, College Park, MD
+1 (240) 413-8060
krezaei@umd.edu
Homepage

EDUCATION

- 2022 – NOW **Doctor of Philosophy**
Computer Science
University of Maryland, College Park
Supervised by Prof. Feizi and Prof. Hajiaghayi
- 2022 – 2024 **Master of Science**
Computer Science
University of Maryland, College Park
- 2018 – 2022 **Bachelor of Science**
RANK 1ST
Computer Engineering
Sharif University of Technology

RESEARCH INTEREST

GenAI interpretability unlearning vision-language models

RESEARCH EXPERIENCE

SEP 2022 – NOW

Graduate Research Assistant, **Reliable AI Lab**, University of Maryland

Research on interpretability, robustness, failure modes, and biases in vision-language generative and discriminative models.

supervised by Prof. Soheil Feizi

FEBRUARY 2025 – APRIL 2025

Student Researcher, **OMEGA**, Google AI

Research on data selection for the task of LLM pretraining; developing a classifier to discard bad tokens.

supervised by Dr. Anton Tsitsulin

MAY 2024 – OCTOBER 2024

Research Intern, **MOSAIC**, Allen Institute for AI (AI2)

Research on large language models unlearning; proposing a new benchmark for *restorative* unlearning.

supervised by Dr. Abhilasha Ravichander and Dr. Faeze Brahman

JULY 2021 – SEPTEMBER 2021

Research Intern, **Theory of Machine Learning Lab**, EPFL

Research on the convergence rate of the optimization algorithms in machine learning.

supervised by Prof. Nicolas Flammarion

JULY 2020 – SEPTEMBER 2020

Research Intern, **Max Planck Institute**

Research on solving semi-definite programming based on Physarum mathematical model.

supervised by Prof. Andreas Karrenbauer

TEACHING EXPERIENCE

Teaching Assistant

FALL 2022 *Design and Analysis of Algorithms (UMD)*

AWARDS

International Collegiate Programming Contest (ICPC)

- 2023 ICPC World Finalist (unable to attend due to visa issues)
- 2023 Ranked 3rd, ICPC North America Championship
- 2020 Ranked 33rd, ICPC World Finals
- 2019 Ranked 1st, Asian Regional Contest
- 2018 Ranked 2nd, Asian Regional Contest

Olympiad in Informatics

- 2018 Silver Medal, 30th International Olympiad in Informatics
- 2018 Silver Medal, 11th Asia-Pacific Informatics Olympiad
- 2017 Gold Medal, 27th Iranian National Olympiad in Informatics
- 2016 Silver Medal, 26th Iranian National Olympiad in Informatics

Awarded the University of Maryland **Dean's Fellowship**

PUBLICATIONS

🔖 PRIME: Prioritizing Interpretability in Failure Mode Extraction

Keivan Rezaei*, Mehrdad Saberi*, Mazda Moayeri, Soheil Feizi

ICLR 2024

🔖 RESTOR: Knowledge Recovery through Machine Unlearning

Keivan Rezaei, Khyathi Chandu, Soheil Feizi, Yejin Choi, Faeze Brahman, Abhilasha Ravichander

ARXIV PREPRINT

🔖 On Mechanistic Knowledge Localization in Text-To-Image Generative Models

S Basu*, **Keivan Rezaei***, P Kattakinda, VI Morariu, N Zhao, RA Rossi, V Manjunatha, S Feizi

ICML 2024

🔖 Text-To-Concept (and Back) via Cross-Model Alignment

Mazda Moayeri*, **Keivan Rezaei***, Maziar Sanjabi, Soheil Feizi

ICML 2023

🔖 Understanding and Mitigating Compositional Issues in Text-to-Image Generative Models

A Zarei*, **K Rezaei***, S Basu, M Saberi, M Moayeri, P Kattakinda, S Feizi

ARXIV PREPRINT

🔖 Run-Off Election: Improved Provable Defense against Data Poisoning Attacks

Keivan Rezaei*, Kiarash Banihashem*, Atoosa Chegini, Soheil Feizi

ICML 2023

🔖 Robustness of AI-Image Detectors: Fundamental Constraints and Practical Attacks

Mehrdad Saberi, Vinu Sankar Sadasivan, **Keivan Rezaei**, Aounon Kumar, Atoosa Chegini, Wenxiao Wang, Soheil Feizi

ICLR 2024

🔖 Multi-agent Delegated Search

α, β Mohammad Hajiaghayi, **Keivan Rezaei**, Suho Shin

EC 2023

FALL 2021 *Design of Algorithms (Sharif)*

FALL 2020 *Compiler Design (Sharif)*

Algorithms and Programming Instructor

2017-2019 *Allame Helli Highschool*

OTHER EXPERIENCES

FALL 2019 - SUMMER 2022

Scientific Member, Iranian National Olympiad in Informatics Committee

I was Iran's team leader at the International Olympiad in Informatics 2021, with additional responsibilities including designing exams to select talented students and conducting preparation camps.

WINTER 2022 - SUMMER 2022

Data Scientist, Torob

Torob is a search engine that allows customers to find the cheapest price for items. I worked on improving, and analyzing Torob's search engine results along with providing some baseline solutions for Torob Challenge.

FALL 2019, FALL 2020

Technical Staff, Sharif AI Challenge

Developed server code and simulated the game in Fall 2019, and designed the game in Fall 2020.

SPRING 2019

Head of Quera College Data Structure Course, Quera

Responsible for providing educational materials.

SKILLS

Machine Learning Libraries, torch, numpy, pandas.

Linux, shell and Bash scripts.

Algorithms and Competitive Programming, as IOI and ICPC awards suggest.

Programming Languages, C++, Python, and Java.

Software Engineering, object-oriented design patterns and other development methodologies.

☐ A Regret Analysis of Repeated Delegated Choice

Mohammad Hajiaghayi, Mohammad Mahdavi, **Keivan Rezaei**, Suho Shin

AAAI 2024

☐ Ad Auctions for LLMs via Retrieval Augmented Generation

α, β Mohammad Hajiaghayi, Sébastien Lahaie, **Keivan Rezaei**, Suho Shin

NEURIPS 2024