Keivan Rezaei

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 | Knowledge Localization | Model Editing | Unlearning | Data Selection for Pretraining | Econ + AI

RESEARCH EXPERIENCE

SEP 2022 – NOW

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

interpretability of generative AI from both **model** and **data** perspectives: **localizing knowledge**, detecting and **explaining failure modes**, and analyzing the influence of individual data points in **unlearning** and **data selection** for pretraining.

FEBRUARY 2025 – APRIL 2025

Student Researcher, **Google Research**

developed new techniques to filter out data in **LLM pretraining**

Mentors: Anton Tsitsulin, Peilin Zhong, and Vahab Mirrokni

MAY 2024 – OCTOBER 2024

Research Intern, **Allen Institute for AI (Ai2)**

proposed a new benchmark for **machine unlearning**, evaluating *restorative* ability of unlearning algorithms

Mentors: Abhilasha Ravichander, Faeze Brahman, and Yejin Choi

MAY 2025 – AUGUST 2025

Research Intern, **Document Intelligence, Adobe**

Proposed a framework for retrieval-augmented generation to visualize scientific designs

Mentors: Ani Nenkova

JULY 2021 – SEPTEMBER 2021

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

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

Mentor: Nicolas Flammarion

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

Machine unlearning

Model State Arithmetic for Machine Unlearning
Keivan Rezaei*, Mehrdad Saberi*, Abhilasha Ravichander, Soheil Feizi
ARXIV PREPRINT

RESTOR: Knowledge Recovery in Machine Unlearning
Keivan Rezaei, Khyathi Chandu, Soheil Feizi, Yejin Choi, Faeze Brahman, Abhilasha Ravichander
TMLR 2025

Interpretability

PRIME: Prioritizing Interpretability in Failure Mode Extraction
Keivan Rezaei*, Mehrdad Saberi*, Mazda Moayeri, Soheil Feizi
ICLR 2024

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

Localizing Knowledge in Diffusion Transformers
Arman Zarei, Samyadeep Basu, **Keivan Rezaei**, Zihao Lin, Sayan Nag, Soheil Feizi
NEURIPS 2025

Understanding and Mitigating Compositional Issues in Text-to-Image Generative Models
S Zarei*, **Keivan Rezaei***, S Basu, M Saberi, M Moayeri, P Kattakinda, S Feizi
ARXIV PREPRINT

TEACHING EXPERIENCE

Teaching Assistant	FALL 2022	*Design and Analysis of Algorithms (UMD)*
	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 served as **Iran's team leader** at the International Olympiad in Informatics 2021, and was responsible for designing exams and conducting training camps to prepare talented students.

WINTER 2022 - SUMMER 2022

Data Scientist, **Torob**

Torob is a search engine that helps customers find the lowest prices for products. I worked on improving and analyzing Torob's search engine results, and also developed baseline solutions for the 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**

Provided online coursework in C++ and algorithmic problem solving.

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.

Economics + AI

Ad Auctions for LLMs via Retrieval Augmented Generation
$^{\alpha,\beta}$ Mohammad Hajiaghayi, Sébastien Lahaie, **Keivan Rezaei**, Suho Shin

NEURIPS 2024

Online Advertisements with LLMs: Opportunities and Challenges
$^{\alpha,\beta}$ Soheil Feizi, MohammadTaghi Hajiaghayi, **Keivan Rezaei**, Suho Shin

ACM SIGECOM EXCHANGES 25

Multi-agent Delegated Search
$^{\alpha,\beta}$ Mohammad Hajiaghayi, **Keivan Rezaei**, Suho Shin

EC 2023

A Regret Analysis of Repeated Delegated Choice
$^{\alpha,\beta}$ Mohammad Hajiaghayi, Mohammad Mahdavi, **Keivan Rezaei**, Suho Shin

AAAI 2024

Robustness

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