# Pardis Sadat Zahraei

✉ paradisez2001@gmail.com    🎓 Google Scholar    in Linkedin    ⌨ Github    🌐 Website

## Education

**B.Sc In Computer Engineering**                                   Nov 2020 - Ongoing
*Sharif University of Technology*                                      *Tehran, Iran*
FinalGrade: 18.4/20 (GPA: 3.87/4)

**Diploma in mathematics and physics**                                      Aug 2020
*Aboureihan HighSchool*                                               *Tehran, Iran*

## Publications

**WSC+: Enhancing The Winograd Schema Challenge Using Tree-of-Experts**
*Authors:* **Pardis Sadat Zahraei**, *Ali Emami*
*The 2024 Conference of the European Chapter of the Association for Computational Linguistics (EACL)*

**TuringQ: Benchmarking AI Comprehension in Theory of Computation**
*Authors:* **Pardis Sadat Zahraei**, *Ehsaneddin Asgari*
*The 2024 Conference on Empirical Methods in Natural Language Processing (EMNLP)*

**Detecting Bias and Enhancing Diagnostic Accuracy in Large Language Models for Healthcare**
*Authors:* **Pardis Sadat Zahraei**, *Zahra Shakeri*
*UnderReview- Nature npj Digital Medicine journal*

**Translate With Care: Addressing Gender Bias, Neutrality, and Reasoning in LLM Translations**
*Authors:* **Pardis Sadat Zahraei**, *Ali Emami*
*UnderReview- The 2025 International Conference on Computational Linguistics (COLING)*

## Research Interests

- RAG and Factual Assessments
- Resources and Evaluation
- Multilingualism and Machine Translation
- LLM Reasoning and Problem-Solving
- Ethics, Bias, and Fairness
- AI for Healthcare

## Research Experience

**Remote Research Project**                                    July 2024 – Present
*University of Maryland College Park*                                        *USA*
Working under the supervision of Professor Jordan Boyd-Graber at the University of Maryland, I focus on the timely aspects of QA systems.

........................................................................................................

**Remote Research Project**                                    July 2024 – Present
*University of Toronto*                                                   *Canada*
I am a research assistant in the Health Informatics, Visualization, and Equity (HIVE) Lab at the University of Toronto, under the supervision of Professor Zahra Shakeri. Our project has resulted in a paper aimed at mitigating bias, evaluating, and benchmarking LLMs, particularly in clinical health. We also conducted an in-depth analysis of demographic factors influencing the models' decision-making abilities. This project was funded with the DSI grant.

........................................................................................................

**Remote Research Project**                                 July 2023 – July 2024
*Brock University*                                                       *Canada*
As a research assistant to Professor Ali Emami, I contributed to research that evaluated and benchmarked the reasoning capabilities of LLMs, uncovered hidden challenges, and explored multilinguality and ethical considerations in translation models for low- to mid-resource languages. This work led to two papers.

........................................................................................................

**Remote Research Project**                                           July 2023 – Present
*Qatar Computing Research Institute*                                           *Qatar*

Under the supervision of Dr. Ehsaneddin Asgari, I contributed to the development of TuringQ, the first benchmark designed to evaluate LLMs' reasoning in computation theory. We explored the concept of using LLMs as judges and investigated the application and importance of understanding computation theory in LLMs. Additionally, I am involved in two ongoing research projects: a survey on generative AI and an exploration of the cultural understanding aspects of LLMs.

## Honors & Awards

**DSI Research Grant Awardee, 2024**
*Awarded CDN $10,000 DSI Critical Investigation of Data Science (CIDS) Research Grant.*

**Top Performer in Machine Learning Course, 2023**
*Secured 1st place in all Sharif University's Kaggle Machine Learning competitions.*

**Full Tuition Scholarship Award, 2020**
*Received full scholarship from Sharif University of Technology for B.Sc. degree.*

**Distinguished Academic Ranking, 2020**
*Ranked in the top 0.3% among 155,000 participants in the Iranian National Mathematics University Entrance Exam.*

## Teaching Experiences

**Teaching assistant**
*Sharif University of Science and Technology*                                           *Tehran, Iran*

- Natural Language Processing | Dr. Asgari                                Spring 2024-2025
- Machine Learning | Dr. Motahari                                Spring 2024-2025
- Machine Learning | Dr. Seyed-Salehi                                Spring 2024-2025
- Machine Learning | Dr. Motahari                                Fall 2023-2024
- Machine Learning | Dr. Jafari                                Fall 2023-2024
- Artificial Intelligence | Dr. Rohban                                Spring 2024-2025
- Artificial Intelligence | Dr. Rohban & Dr. Soleymani                                Fall 2023-2024
- Artificial Intelligence | Dr. Samiei                                Spring 2022-2023
- Advanced Information Retrieval | Dr. Soleymani                                Spring 2024-2025
- Linear Algebra | Dr. Sharifi (Head TA)                                Fall 2023-2024
- Linear Algebra | Dr. Rabiee                                Spring 2022-2023
- Engineering Probability and Statistics | Dr. Najafi                                Fall 2023-2024
- Engineering Probability and Statistics | Dr. Sharifi                                Spring 2022-2023
- Compiler Design | Dr. Hosseinmardi                                Fall 2023-2024
- Compiler Design | Dr. Hosseinmardi                                Spring 2022-2023
- Numerical Computation | Dr. Hossein-Ghorban                                Fall 2023-2024
- Numerical Computation | Dr. Hossein-Ghorban                                Spring 2022-2023

## Relevant Courses

**Generative Models (Graduate Level)** | ONGOING
**Natural Language Processing (Graduate Level)** | 20/20
**Machine Learning** | 20/20
**Artificial Intelligence** | 19.8/20
**Game Theory** | 20/20
**Numerical Computation** | 19.6/20
**Data Structures and Algorithms** | 20/20
**Linear Algebra** | 20/20
**Engineering Probability and Statistics** | 19.4/20
**Generative AI with Large Language Models** Certificate
**Data Visualization** Certificate
**Text Retrieval and Search Engines** Certificate
**Text Mining** Certificate

## Projects

### Advanced Information Retrieval Course Project ⍾ Github
*Built a search engine from scratch covering key IR concepts such as exploratory data analysis, index construction/compression, word embeddings, clustering techniques (hierarchical, DBSCAN, K-means), generative models, and web crawling.*

### Machine Learning Course Project ⍾ Github
*Applied various machine learning techniques, including Decision Trees, Regression, Naïve Bayes, SVM, Neural Networks, LSTM+CNN, and Ensemble Methods.*

### Sentiment Analysis on Twitter and YouTube Data ⍾ Github
*Performed sentiment analysis using a custom preprocessing pipeline involving lemmatization, tokenization, Named Entity Recognition, spell checking, bigram checking, contradiction fixing, and a custom slang dictionary.*

### Cross-Lingual Drug Name Prediction ⍾ Github
*Used embedding methods (FastText and BERT) to predict appropriate drug names in both Persian and English languages.*

### Persian Summarization and NER ⊕ NER Model ⊕ Summarizer
*Developed a token classification system using transformer-based models for Named Entity Recognition and created another model for generating summaries in Persian.*

### Persian Formal-Informal Text Translation ⊕ PersianTextFormalizer ⊕ PersianEase
*Fine-tuned Persian language models to generate formal text from informal inputs and vice versa.*

## Languages

- English (Advanced)
  (TOEFL iBT: 113/120, R:28, L:28, S:28, W:29))
- Persian (Native)

## NLP Skills

- Python (Transformers, Torch, Scikit-learn)
- Prompt Engineering
- Fine-tuning (SFT, TRL, DPO)
- Web Crawling & RAG Systems
- PEFT & LoRA
- Distillation & Quantization

## Extracurricular Activities

### EMNLP 2024 Student Volunteer
Assisting EMNLP 2024 AV Technician Team in managing various online content tasks.

### Sharif University Large Language Models Journal Club
Reviewed and discussed research papers on large language models weekly.

### Sharif University Winter Seminar Series on Advanced Topics in Computer Science
Served on the executive team for the 8th and 9th Winter Seminar Series.

### Sharif University Environmental Enthusiasts Scientific Association
Participated in environmental awareness events.