

Maryam Yazdi

🌐 <https://maryamyazdi.github.io/> | 🐙 github.com/maryamyazdi | ✉ maryam.yazzdi@gmail.com | 📍 Iran

RESEARCH INTERESTS

- Large Language Models (robustness, various application domains)
- Knowledge Extraction
- Trustworthy Machine Learning

EDUCATION

Bachelor of Science, Computer Software Engineering

University of Isfahan

Aug. 2018 - May 2023

Isfahan, Iran

EXPERIENCE

Data Scientist

DataCoLab

May 2023 - June 2024

London, UK (Remote)

- Implemented over 10 different LLM-powered **AI agents** to satisfy complex user queries and support seamless collaborative task execution over structured or unstructured data.
- **Fine-tuned** and deployed open-source LLMs to effectively replace GPT-based inference pipelines, reducing the company's OpenAI API costs by 30%, and enhancing system scalability and operational autonomy.
- Conducted **deep-dive analysis** on available datasets to design innovative data-driven solutions addressing specific requirements of the projects.

Teaching Assistant

University of Isfahan - Computer Science Department

Aug. 2021 - June 2022

Isfahan, Iran

- Provided support for Data Mining and Principles of Programming courses.
- Conducted detailed grading and provided personalized feedback for assignments and exams.
- Held weekly sessions to assist students with course material and answer questions about assignments.

SKILLS

- **Technologies & Frameworks:** Python, Pytorch, Transformers, Scikit-learn, PEFT, Pandas, NumPy, SQL, Git/Github, LaTeX
- **English Proficiency:** Fluent - IELTS Academic Overall Band Score 7.5

PROJECTS

RAG-based Question Answering Bot

April 2024

- [Github repository](#)
- Implemented a custom retrieval-augmented generation (RAG) framework from scratch, integrated with LLM-based question answering and a convenient user interface.
- Technologies: FaissDB, LLMs, Docker, Flask, streamlit

Paper Summarizer Agent

July 2023

- Engineered chained-prompts with different summarization pipelines, enabling comprehensive research paper condensation across diverse LLMs regardless of context window size.
- Technologies: Python, Langchain, LLMs, streamlit

Contextual Machine Translation for Movies Subtitles

June 2023

Advisor: Prof. Reza Ramezani

- Enhanced training data by extracting and validating English-Persian idiom translations from movie subtitles.
- Fine-tuned a language model for translating English movie subtitles to Persian, emphasizing idioms and slang.
- Technologies: Pytorch, Beautiful Soup, Transformers