

Maryam Yazdi

github.com/maryamyazdi | maryam.yazzdi@gmail.com | linkedin.com/in/maryamyazdi | Isfahan, Iran

EDUCATION

Bachelor of Science, Computer Science
University of Isfahan

Sep. 2018 - May 2023
Isfahan, Iran

EXPERIENCE

Data Scientist

DataCoLab

May 2023 - June 2024

London, UK (Remote)

- Developed a coherent platform consisting of over 10 different LLM-powered **AI agents** to satisfy complex user queries through decision-making and autonomous/collaborative task execution.
- Replaced a GPT-outsourced task with a **fine-tuned**, locally hosted open-source language model, significantly reducing the company's costs for OpenAI API.
- Conducted **deep-dive analysis** on provided datasets to design data-driven solutions addressing specific requirements of the project.

Teaching Assistant

University of Isfahan - Computer Science Department

May 2020 - July 2021

Isfahan, Iran

- Courses: Data Mining (Dr. Reza Ramezani), Information Retrieval (Dr. Marzieh Hosseini)
- Conducted detailed grading and provided personalized feedback for assignments and exams.
- Held weekly office hours to assist students with course material and answer questions about assignments.

TECHNICAL SKILLS

- **Languages:** Python, SQL, C#
- **Technologies/Frameworks:** Pytorch, Huggingface Transformers, PEFT, Keras, Pandas, NumPy, Sci-Kit Learn, Plotly, Flask, Docker, Unit testing, Git/Github
- **Areas of Interest:** Large Language Models and Their Applications, Natural Language Processing, Information Retrieval

PROJECTS

RAG-based Question Answering Bot

April 2024

- [Github repository](#)
- Developed a bot to answer user questions using the selected language model, supported by the context provided by a custom retrieval-augmented generation (RAG) framework implemented from scratch.
- Technologies: FaissDB, Docker, Google/OpenAI models, streamlit

Paper Summarizer Agent

July 2023

- Implemented a chain-like summarization pipeline, compatible with diverse LLMs with different window-contexts to authentically summarize a research paper.
- Technologies: Python, Langchain, Google/OpenAI models, streamlit

Contextual Machine Translation for Movie Subtitles

June 2023

Guide: Prof. Reza Ramezani

- Improved the limited available training data by utilizing language models to validate new data, extracted from top-rated Persian translations of English movie subtitles.
- Fine-tuned a language model with a focus on translating English-Persian idioms and slang used in movie subtitles.