

MAHDI KHOSHARAMZADE

Tehran, Iran

☎ (+98)9147716617 ✉ Mahdikhoshmaramzade@gmail.com  [Github](#)  [Personal WebPage](#)

Education

M.S. of Software Engineering - K.N.Toosi University of Technology 2022 – ongoing

Thesis: Improving relation extraction models in medical domain using large language models

GPA: 3.65/4

Class Rank: Among top 10%

B.S. of Electrical Engineering - University of Tabriz 2017 – 2021

Thesis: Handwritten digit recognition using ResNet architecture

Major GPA: 3.34/4

Research Interests

I am passionate about improving the reliability, interpretability, and robustness of large language models. My work focuses on tackling issues like hallucinations and biases while enhancing overall model performance. I aim to develop methods that make these models more transparent, trustworthy, and aligned with human values, ensuring they perform effectively in real-world applications.

Currently, My master's thesis focuses on converting unstructured data from electronic health records into accessible tables using LLMs and RAG framework, potentially revolutionizing clinical research and healthcare outcomes. Further, under the guidance of my supervisor, I contributed to an industrial project in Italy aimed at enhancing medical question-answering systems. This project focused on reducing hallucinations and improving the reasoning capabilities of LLMs using RAG systems and knowledge graphs. I led the optimization of the retrieval pipeline to extract relevant sub-graphs from extensive medical knowledge graphs based on user inputs. I helped improve retrieval accuracy and efficiency by employing advanced community detection graph algorithms such as the Louvain algorithm. In the final phase, I designed a chain of thought prompts to make the model more transparent.

Publications

- **Mahdi Khoshmaramzade**, Saeed Farzi, Sina Mansoori *Working Paper*
Efficient Medical Relation Extraction via RAG-Enhanced LLMs and Knowledge Graphs
In This paper, our focus is to enhance model's reliability and reasoning and Diminishing it's Hallucination using Custom Large Knowledge graph in the medical Domain
- Boshra Pishgoo, **Mahdi Khoshmaramzade** *Working paper*
Clustering and behavioral analysis of social media users using Large Language models
This paper aims to extract types of human behaviors by gathering data from social media platforms such as X. The study employs large language models (LLMs) and clustering algorithms, including DBSCAN and HDBSCAN, to accomplish this task.

Research Experience

K.N.Toosi University of Technology, Research assistant Jun. 2023 – Ongoing

- We are working on LLMs systems and RAG pipeline especially retrieval part to extract optimized sub-graphs from large medical knowledge graph, based on user's input to LLM
- We have worked on Several graph algorithms such as Community detection to improve retrieval part

Beshart, NLP Researcher Oct. 2022 – Mar. 2023

- Fine-tuned BERT model to analyze Twitter data, identifying unmet needs of Indian consumers.
- Extracted insights on gaps in the Indian market by processing and analyzing social media data.

Projects

🔗 RAG-driven Llama3.1 using UMLS & Knowledge graphs for Relation extracion Neo4j, Langchain	ongoing
🔗 Dialogue-Summarization Using T5 Language model HuggingFace Transformer, Pytorch, T5	2025
🔗 Few-Shot Chemical-Disease Relation(CDR) Extraction on PubMed Abstracts Langchain, Python	2024
🔗 Medical QA LLM model enhanced By RAG and KG Langchain, Neo4j, LLaMA 3, UMLS(NIH), Python	2024
🔗 Develop a pure python script to crawl Instagram pages Python	2024
🔗 Masked Language Model Implementation on Middle Persian language Pytorch, HuggingFace	2023
🔗 Representation learning using Deepwalk algorithm on MovieLens dataset Python	2023
🔗 Build and train Question Classification models using TREC dataset python, Scikit Learn, SQL Servers	2022
🔗 Emotion Recognition by Textual Tweets Classification Using Voting Classifier(LR-SGD) Python, NLTK	2022
🔗 Search engine for Persian Poems Python, whoosh, Hazm, Parsivar	2022
🔗 Designing a software architecture for web-based chess game	2021

Teaching Assistant Experiences

Algorithm Design - K.N.Toosi University of Technology	Ongoing
Role: Head TA in the course	Dr. Pishgoo
Artificial intelligence - K.N.Toosi University of Technology	Fall 2023
Designing HWs and Teaching DL Frameworks and libraries such as PyTorch as a Head Teaching Assistant.	Dr. Pishgoo
Advanced Programming - Sharif University of Technology	Fall 2022
First-class Functions, Closures and Decorators and OOP in Python were taught.	Dr. Sharifi-Zarchi

Technical Skills

Programming Languages: Python, Rust, SQL, Cypher, L^AT_EX
NLP: LangChain, LangGraph, Hugging Face Tokenizers and Transformers, NLTK, Ollama, Neo4j, OpenAI
ML: PyTorch, Matplotlib, Numpy, Pandas, Scikit Learn
Web-development: Django, HTML, CSS, JavaScript, FastAPI
Data Management: Neo4j, MySQL, Microsoft SQL Server
Other Tools & Skills: Microsoft Power BI, Git, Docker, Linux, TeXstudio, Prompt Engineering, Wireshark

Certificates

🔗 Generative AI with Large Language Models DeepLearning.AI, AWS	2024
🔗 Neural Networks and Deep Learning DeepLearning.AI	2023
🔗 Supervised Machine Learning: Regression and Classification DeepLearning.AI	2023
🔗 Mathematics for Machine Learning: Multivariate Calculus Coursera	2021
🔗 Mathematitcs for machine learning: Linear Algebra Coursera	2021
🔗 Capstone: Retrieving, Processing, and Visualizing Data with Python Coursera	2021
🔗 Using Databases with Python Coursera	2021
🔗 Using Python to Access Web Data Coursera	2021
🔗 Python Data Structures Coursera	2021

Languages

English: Fluent (TOEFL score: 82)
Persian: Native
Azari: Native
Turkish: Fluent

❗ References, Further information, and Proofs are available upon Request