Erfan Shahabi

AI Software Engineer

+98 912 640 40 96 | ErfanShahabi@outlook.com | ErfanShahabi.github.io

Linkedin | GitHub | Tehran, Iran

EDUCATION

University Of Tehran   Sep 2024 - Present

M.Sc in Information Technology Engineering   Tehran, Iran

Supervisor: Dr. Azadeh Shakery

Islamic Azad University, Science and Research Branch   Sep 2019 - Aug 2023

B.Sc in Industrial Engineering, GPA: 4/4 (18.07/20)   Tehran, Iran

Supervisor: Dr. Abbas Saghaei

HIGHLIGHTED PROJECTS

NLP :

- **NER and Rule-Based Text Processing [🌐]**
- **Multiview Representation Learning and Evaluation in NLP: Classical IR, Word Embeddings, and Vision-Language Models [🌐]**
- **Deep Learning for Sequence Labeling and Text Summarization with RNNs and Attention [🌐]**
- **Multilingual NER and Efficient Summarization with Large Language Models and Model Compression [🌐]**
- **Evaluating and Prompt Engineering for Machine Translation using LLM APIs [🌐]**
- **ChatFood: Building a ReAct-based Conversational Food Agent with LangGraph and Gemini [🌐]**

LLMs:

- **Prompt Engineering and Output Evaluation for LLMs on Persian Question Answering [🌐]**
- **Fine-Tuning LLaMA-3.2 using ORPO for Preference-Based Alignment [🌐]**
- **Alignment of Large Language Models via Human Preference Optimization and Direct Preference Objective (DPO) [🌐]**
- **Quantization-Aware Reasoning and Text-to-SQL Agent Design using ReAct and LangGraph [🌐]**
- **LLM-Based Multi-Agent User Simulation for Evaluating Conversational Recommender Systems [🌐]**

Intelligent Information Retrieval:

- **Document Ranking and Evaluation using Inverted Index and BM25 Retrieval [🌐]**
- **Hyperparameter Sensitivity Analysis of BM25 in Large-Scale Document Retrieval [🌐]**
- **Scalable Link Analysis and Opinion Mining using MapReduce, HITS, and PageRank [🌐]**
- **Sentiment Classification of IMDB Reviews using Text Preprocessing and Classical IR Techniques [🌐]**
- **Relevance Classification using BERT for Query-Document Pairs in IR Tasks [🌐]**
- **Conversational Search with Retrieval-Augmented Generation: Dataset Construction, RAG Design, and Multi-faceted Evaluation [🌐]**

Machine Learning and Data Analysis:

- **Multi-City Real Estate Data Cleaning and Integration for Exploratory Data Mining [🌐]**
- **Designing a Star Schema Data Warehouse for Supermarket Sales Analysis [🌐]**
- **Market Basket Analysis and Association Rule Mining from Online Retail Transactions [🌐]**
- **Interactive Comparison of Classification Models using Decision Boundary Visualization [🌐]**
- **Semantic Clustering of StackOverflow Questions using Sentence Embeddings and Dimensionality Reduction [🌐]**
- **Forecasting Community Dynamics in Social Graphs using Multi-Task Graph Neural Networks [🌐]**

Bachelor Thesis:

- **Used Car Price Prediction Based on Manufacturing Year and Mileage Using Machine Learning Methods**

EXPERIENCES AND COURSES

Natural Language Processing

University Of Tehran - Supervisor: Dr. Hesham Faili

- Progressed from fundamental NLP concepts to advanced techniques
- Implemented a chatbot as the final project using tool-use strategies
- Utilized APIs, Transformers, Hugging Face, Langchain, Langgraph, and agents

Large Language Models

University Of Tehran - Supervisors: Dr. MJ Dousti, Dr. Y Yaghoob Zadeh

- Gained strong skills working with large language models (LLMs) through multiple practical projects
- Worked extensively with APIs, Transformers, Hugging Face, Langchain, Langgraph, and agents
- Developed use cases showcasing model integration and API utilization

Intelligent Information Retrieval

University Of Tehran - Supervisor: Dr. Azadeh Shakery

- Started with ad-hoc retrieval methods and embeddings
- Worked with BERT models for improved retrieval performance
- Explored the application of LLMs in retrieval-augmented generation (RAG)

Advanced Data Mining

University Of Tehran - Supervisor: Dr. Azadeh Shakery

- Learned data fundamentals including cleaning and preprocessing
- Covered advanced topics such as data cubes, classification, clustering, and deep learning applications
- Applied techniques in real-world datasets

Social Networks

University Of Tehran - Supervisor: Dr. Masoud Asadpour

- Studied graph theory, centrality measures, and community detection
- Applied concepts to analyze real-world network data through projects

TECHNICAL SKILLS

- **Soft Skills:** Teamwork, Meeting Tight Deadlines
- **Programming Languages:** Python, Javascript
- **Large Language Models:** Hugging Face, LangChain, Lang graph, LangFuse, Unsloth, PEFT, Datasets, TRL
- **Databases:** ElasticSearch, Milvus, lanceDB, Redis, PostgreSQL
- **Machine Learning:** PyTorch, Scikit Learn, evaluate, TensorFlow
- **Retrieval:** Sentence Transformers, Lexical Retrieval, Semantic Retrieval, ColBERT, BM25

HONORS AND AWARDS

- **Ranked 14 in Nationwide Graduate Level University Entrance Test of Information Technology Engineering**