

Amirhossein Layegh

KTH Royal Institute of Technology
Stockholm, Sweden
+46 (73) 49 13 530
amlk@kth.se
amirlayegh.github.io
amirhosseinlayegh



Research Interests

Scalable Machine Learning (ML) and Deep Learning (DL), Natural Language Processing, Information Extraction, Distributed Systems

Education

- 2021 – now **PhD in Information and Communication Technology**, KTH Royal Institute of Technology, Sweden
Thesis Title: Information Extraction using Large Language Models.
Advisers: Prof. Mihhail Matskin and Dr. Amir H. Payberah
Focused on advancing methodologies for extracting structured information from unstructured text using the capabilities of large language models, aiming to contribute novel approaches to the field.
- 2018 – 2019 **MSc in Big Data Science**, Queen Mary University of London (QMUL), UK
Thesis Title: Implementation of a Recommendation System for a retail store based on basket analysis.
Adviser: Dr. Arman Khouzani
Developed a recommendation system for Koolbitz, a London-based retailer, using the Apriori algorithm for basket analysis. The system identifies patterns in purchase data to suggest products to customers, enhancing sales and customer experience.
- 2012 – 2017 **BSc in Software Engineering**, Ferdowsi University of Mashhad (FUM), Iran
Adviser: Dr. Mohsen Kahani
Completed a final project on developing a database of academic publications for current and former Ferdowsi University affiliates. Utilized web scraping techniques to automate publication data collection, significantly enriching the university's academic resources.

Industrial Experience

- 2020 – 2021 **Data Scientist**, Neshan Maps, Iran
Employed advanced speed estimate techniques on large-scale user datasets to analyze and determine traffic flow for individual road segments. Developed models to accurately forecast total traffic flow, enhancing the app's navigation accuracy and user satisfaction.
- 2019 – 2020 **Data Analyst**, Koolbitz Ltd., UK
Developed and implemented machine learning algorithms to analyze large sales datasets, identifying key patterns and trends for product recommendations. This significantly improved targeted marketing strategies and customer engagement.
- 2017 – 2018 **Software Developer**, Ranir, Iran
Engaged as an Oracle Application Development Framework (ADF) developer, focusing on building Java-based enterprise applications.

Teaching and Supervision

Supervision

- MSc thesis supervisor of Master students
 - **Deciding Agents Behaviours Through LLMs:** This thesis investigates using LLMs for decision-making and behavior control in AI agents within gaming virtual environments. It aims to enhance AI agent design by integrating natural language processing for intuitive human-AI interactions with potential gaming, virtual reality, and digital assistant applications.
 - **Enhancing Breast Cancer Information Retrieval:** This project focuses on developing, evaluating, and optimizing Retrieval-Augmented Large Language Models specifically for improving breast cancer information retrieval accuracy and efficiency. It aims to contribute to the medical field by providing better tools for accessing and understanding breast cancer-related data.
 - **Designing Novel Effective Methods for Large Language Model Compression:** Guiding research on pruning and lightweight techniques for compressing large language models, specifically targeting 7B - 13B models. The objective is to reduce these models to fewer than 3B parameters while preserving their performance, employing strategies such as knowledge distillation through teacher-student models.

Teaching

- Modern Methods in Software Engineering, KTH, 2021–now
- Distributed AI and Intelligent Agents, KTH, 2021–now
- Databases, FUM, 2016

Publications

- REA: Refine-Estimate-Answer Prompting for Zero-Shot Relation Extraction, Submitted to NLDB 2024
- Wiki-based Prompts for Enhancing Relation Extraction using Language Models, Accepted as full paper at SAC 2024
- ContrastNER: Contrastive-based Prompt Tuning for Few-shot NER (2023), presented as full paper at COMPSAC 2023
- Datacloudsl: Textual and Visual Presentation of Big Data Pipelines (2022)
- A survey of big data pipeline orchestration tools from the perspective of the datacloud project (2021)