

# Mona Pouresmaeil

## MSc Computer Engineering – AI and Data Analytics

Turin, Piedmont, Italy



✉ mona.pouresmaeil@studenti.polito.it | 🌐 Portfolio | 🐙 GitHub | 🔗 LinkedIn | 📊 Kaggle

### Professional Summary

**AI Engineer** specialized in **NLP**, **RAG**, and **agentic workflows** with a proven track record of building production-grade LLM applications. Currently finalizing a Master's thesis at **Brain Technologies** focused on advanced document validation frameworks, with **full immediate availability** for full-time roles.

Co-author of the **SemEval 2026 Task 1** system that ranked **2nd internationally**. Expert in developing fully **custom AI pipelines from scratch** in pure Python, leveraging **LangGraph**, **NLP** techniques, and hybrid search to solve complex industrial problems.

Actively seeking **AI/ML Engineer** or **LLM Engineer** positions to apply advanced expertise in **NLP** and information retrieval systems immediately.

### Education

<b>Politecnico di Torino</b>	<i>Turin, Italy</i>	2023 – July 2026
<ul style="list-style-type: none"><li>MSc in Computer Engineering – <b>Artificial Intelligence &amp; Data Analytics Track</b></li></ul>		
<b>Shariaty Technical and Vocational College</b>	<i>Tehran, Iran</i>	
<ul style="list-style-type: none"><li>BSc in Computer Technology Software Engineering <span style="float: right;">2020 – 2022</span></li><li>Associate Degree in Computer Software Engineering <span style="float: right;">2018 – 2020</span></li></ul>		

### Experience

<b>AI Engineer Intern</b> Brain Technologies	<i>Turin, Italy</i> (On-site)	Jan 2026 – Present
<ul style="list-style-type: none"><li>Finalizing Master's thesis: developed a fully custom AI framework for automatic search, validation, and verification of documentary content using <b>RAG</b>, <b>advanced NLP techniques</b>, and <b>LLM-based pipelines</b> built entirely from scratch</li><li>Designed and implemented the complete RAG pipeline manually, including <b>custom chunking strategies</b>, embedding generation, <b>hybrid search (vector + keyword)</b>, <b>confidence scoring</b> for <b>hallucination mitigation</b>, and compliance features</li><li>Gained hands-on experience building <b>production-oriented LLM applications</b> and <b>information retrieval</b> systems from the ground up in pure <b>Python</b></li></ul>		
<b>Backend Developer Intern</b> Rastava Academy	<i>Tehran, Iran</i> (Remote)	Mar 2020 – Aug 2020
<ul style="list-style-type: none"><li>Developed scalable backend services using <b>Python and Django</b></li><li>Collaborated in <b>Agile</b> teams with <b>Git/GitHub</b> for version control and code reviews</li><li>Designed and optimized <b>database schemas</b> for data integrity and efficient retrieval</li></ul>		
<b>Software Engineering Intern</b> Kasra Company	<i>Tehran, Iran</i> (On-site)	Sep 2019 – Feb 2020
<ul style="list-style-type: none"><li>Customized enterprise software solutions for over 500 clients according to specific business needs and workflows</li><li>Optimized <b>SQL queries</b> to improve database performance and reduce latency</li><li>Participated actively in <b>Agile development cycles</b></li></ul>		

## Key Projects

---

### SemEval 2026 Task 1 – Humor Generation

- Co-developed humor generation system that ranked **2nd internationally**
- Implemented **LLM-based generation** and evaluation workflow leveraging **RAG** and **preference optimization** techniques
- **Tech:** Python, NLP, RAG, **DPO**, **Large Language Models**

### RAG Document Validation Framework

- Built AI system for automatic search, validation, and verification of documentary content using **RAG** and **LangGraph**
- Implemented ingestion for PDF/DOCX/PPTX, local **Llama 3** Q&A via **Ollama**, and answer/requirement verification for compliance workflows
- **Tech:** Python, **LangChain**, **LangGraph**, Ollama (Llama 3), **ChromaDB**, **FAISS**

### Agentic Job Tracker 2026

- Built AI-powered tool for Italian job postings with auto-translation, **field extraction**, tax estimator, and **agentic chat interface**
- Extracted key fields (company, location, RAL, tech stack) and generated **structured summaries** to speed up application triage
- **Tech:** Python 3.13, **Streamlit**, **Llama 3.3-70B**, **Hugging Face**, **Agentic AI**

### Fake News Detection System

- Created complete **NLP pipeline** with **TF-IDF** and classification models (**SVM**, **Logistic Regression**, **LSTM**)
- Compared models and tuned preprocessing/features to improve robustness across datasets
- **Tech:** Python, **scikit-learn**, NLTK, LSTM

### Deep Learning Handwritten Digit Classifier

- Implemented advanced **CNN** in **PyTorch** achieving **99%+ accuracy** on MNIST
- Improved generalization using **batch normalization**, **dropout**, and data augmentation during training
- **Tech:** PyTorch, CNN, Batch Normalization, Dropout

## Skills

---

- **AI & Machine Learning:** **Machine Learning**, **Deep Learning**, **NLP**, **LLMs**, **Retrieval-Augmented Generation (RAG)**, **Agentic AI**, Feature Engineering, **Hallucination Mitigation**, **Prompt Engineering**, Information Retrieval, Hybrid Search, Custom Chunking & Embedding
- **RAG Engineering:** From-scratch pipeline development in pure Python, custom chunking, embedding workflows, hybrid retrieval (vector + keyword), and confidence scoring
- **Frameworks & Libraries:** **Python**, **PyTorch**, **scikit-learn**, **LangChain**, **LangGraph**, **Hugging Face**, TensorFlow, XGBoost, Pandas, NumPy, spaCy, NLTK, **Ragas**, **Qdrant**, **Docling**, **Unsloth**, **ChromaDB**, **FAISS**
- **Software Engineering & Tools:** Django, SQL, Query Optimization, React, **REST APIs**, Git/GitHub, **Docker**, **AWS**, Streamlit, Flask, Agile, Jupyter, Google Colab

## Certifications

---

- **AI Agents Intensive Course with Google** (2025)
- **RAG for Professionals with LangGraph, Python and OpenAI** – Udemy (Dec 2025)
- Introduction to TensorFlow for Artificial Intelligence, Machine Learning, and Deep Learning – DeepLearning.AI (Nov 2025)
- GenAI Intensive Program – Kaggle

## Languages

---

- **English** — C1 (Fluent)
- **Italian** — A2-B1 (actively improving)