Rocco Lo Conte

Contacts: +39 3662359991, Email, LinkedIn, GitHub

Lives in: Avellino, Italy (Open to Relocate: Yes, Nationality: Italian)

Languages: Italian (Mother Tongue), English (CEFR: C1)

Years of Experience: 1.5

Programming Languages: Python, JavaScript, C#, SQL, Java, C

Tools: PyTorch, .NET, Azure, Docker, Jenkins, Git



Software Engineer (AI/ML Focus)

04/24 - Present

Remote

Reply, WM Rome Team

- Context: Fast-paced startup environment (15 people) requiring full-stack ownership: design, development, client communication, costs estimation and production support across multiple concurrent projects.
- Impact: Delivered 8 of 12 concurrent client projects (67% of company portfolio) by collaborating in Agile teams of 3 members each, contributing to €1M in secured contracts.
- AI Development: Co-leading end-to-end development of an AI-powered call screening system processing 3K monthly interactions as one of two AI/ML specialists in the company, reducing human screening costs by automating initial financial eligibility assessments through fine-tuned LLM classification.
- **Production Engineering**: Implemented comprehensive logging systems and resolved critical production issues through systematic debugging.
- RAG: Implemented a RAG system using Azure OpenAI (GPT-40) and Azure AI Search, reducing hallucination rate by 80% through systematic prompt optimization, validated across 25 test cases with human evaluation.
- ETL: Built ETL pipelines processing 5 enterprise databases containing millions of records.
- DevOps: Automated deployment of 5 resources through CI/CD pipelines developed with Jenkins on Azure.

AI Evaluator (Freelance, Weekends Only)

03/25 – Present

Outlier AI Remote

• Conducting part-time AI evaluation tasks (RLHF on coding prompts) on weekends when available.

Software Engineer Intern

02/19 - 05/19

University of the Bundeswehr Munich

Munich, Germany

• Built InputGraph, a Java-based interactive interface for the <u>Grappa project</u>, visualizing parsed data to support analysis in <u>ParVis</u>, a parser behavior tool.

Projects

Online Shortest Path with Machine Learned Advice (Master's Thesis)

2024

• Created an AI algorithm in Python under the supervision of Prof. Flavio Chierichetti that intelligently combines multiple navigation services (TomTom, Google Maps) to provide optimal real-time route recommendations. Tested across 75 locations between Rome and Paris, performing nearly as well (up to 10%) as having perfect knowledge of which service would be fastest for each specific route.

Pre-trained a language model

2025

• Pre-trained a 6-layer OLMO-2-inspired transformer model on the Shakespeare dataset with a custom BPE tokenizer by using PyTorch and personal GPU (MPS), achieving 2.77 perplexity.

Fine-tuned a language model

2025

• Fine-tuned GPT2 by using LoRA and personal GPU (MPS) on 3K samples of the Databricks Dolly-15k, achieving a 3% improvement on ROGUE/BLEU scores.

Trained a vision model

2025

• Trained a ViT on the MNIST dataset by using PyTorch and personal GPU (MPS), achieving 97% test accuracy.

Obtain routing data from 8 REST APIs in Python

2024

EDUCATION

Sapienza University of Rome	09/20 - 01/24
MSc in Computer Science, Language: English, Grade: 107/110, CGPA: 27.5/30.0	Rome, Italy
University of Salerno	09/16 - 05/20
BSc in Computer Science, Language: Italian, Grade: 94/110, CGPA: 24.1/30.0	Salerno, Italy