Vittorio Rossi

rossi.vittori@gmail.com | +39 3458367454 | linkedin.com/in/vitrossi | github.com/VittorioRossi

PROFESSIONAL EXPERIENCE

Commerce Clarity
Data Scientist & Machine Learning Engineer

January 2024 – Today
Italy

- Curated **100K+ e-commerce conversations** for open source model fine-tuning of a SLM
- Developed prompt versioning and logging system to track **10M+ prompts executions** with LLMs
- Designed computer vision pipeline to match **10k products across 5M competitors**
- Trained categorization pipeline from scratch, correctly classifying **10k products within seconds**

Q8 Kuwait Petroleum Italy S.p.A.
Data Scientists Intern

June 2023 – November 2023
Rome, Italy

- **Led a 4-person team** to enhance IoT backend, reducing measurement errors by **50%**
- Supported migration of data platform on AWS, migrated **5+ jobs to AWS Glue**
- Extracted insights from complex dataset of **3M daily entries from 200+ IoT devices**

UniversalDot Foundation
Machine Learning Expert

June 2021 – November 2021
Remote

- Designed an RAG-based recommendation system matching curriculums to gigs (GitHub)

RESERCH, PROJECTS & PUBLICATIONS

More Efficient Text Diffusion via Length Prediction
Vittorio Rossi, Dirk Hovy, et al.

Manuscript in preparation

- Developed **novel inference technique for Diffusion LLMs** using lightweight 1-layer MLP to predict variable sequence length from hidden states at inference time, enabling **significant compute savings** without costly model retraining

LLM Forecasters - Zero-Shot Time Series Forecasting
Advisor: Daniele Tonini

BSc Thesis

- Investigated LLMs as **zero-shot forecasting models** against traditional statistical methods (ARIMA, Prophet) and proposed **comprehensive novel benchmark** to systematically evaluate this emerging capability across multiple datasets

CodeGraph
Full stack, Python, JS, Neo4j, FastAPI

Open Source Project
January 2025

- Developed **full-stack application** with React frontend and FastAPI backend that creates structured knowledge graphs of codebases using Neo4j, specifically designed for **LLM querying** to analyze complex code dependencies

Crowd Flow Estimation
PyTorch, VGG19, ResNet50

Computer Vision Research
January 2025

- Investigated impact of receptive field size on crowd density map estimation using systematic depth variations of VGG19 and ResNet50 encoder-decoder architectures, evaluating performance on ShanghaiTech dataset with comprehensive analysis

EDUCATION

Bocconi University
MSc in Artificial Intelligence

September 2024 – Today
Milan, Italy

- Relevant coursework: NLP, CV, Software Engineering, Deep Learning and RL, Advanced Linear Algebra

Georgia Institute of Technology
Exchange Student

January 2024 – April 2024
Atlanta, USA

- Relevant coursework: Undestanding Marktes with Data Science, Marketing Analytics

Bocconi University
BSc in Economic and Management and Computer Science

September 2021 – July 2024
Milan, Italy

- Final Grade: 110/110
- Relevant coursework: Machine Learning, Computer Programming, Mathematics, Statistics

SKILLS & LANGUAGES

Languages:	Italian (Native speaker), English (Fluent, IELTS 7.5), Spanish (Basics)
Tech skills:	Python, PHP, JavaScript, C/C++, PyTorch, Numpy, Pandas, SQL, Spark, MongoDB, Linux, Cloud, HPC, SLURM, Research methods, software design, project management
Hobbies:	Football, sports, coding, math and physics

I hereby authorize the treatment of my personal data according to the local law