Luca **Di Vita**

Profile

Machine Learning Engineer with a degree in Computer Engineering and a strong background in **Python** and **software engineering**. I design, build, and deploy **production-grade ML systems**, working across the entire lifecycle—from data pipelines to real-time inference and scalable microservices. Curious and research-driven, I also explore **evolutionary computation** and experimental approaches to AI. Active in the tech community as GDG Pescara organizer and Python Pescara co-founder, I frequently deliver technical talks and contribute to training and knowledge-sharing initiatives, aiming to make complex concepts accessible and practical.

Contacts

@ lucadivita.ldv@gmail.com
Git - lucadivit
LinkedIn - lucadivit
lucadivita.it

Personal Information

Citizenship: **Italian**

Languages: **Italian** (native), **English** (professional working proficiency)

Skills

- **Machine Learning & Data**: data processing and feature engineering, model development and evaluation; scikit-learn, PyTorch, Keras, Pandas, NumPy.
- **Backend & MLOps**: Python, Flask, FastAPI, Docker; development of APIs and microservices for ML deployment.
- **Data Engineering**: SQL, data pipelines, ETL workflows.
- **Tools & Dev Practices**: Git, containerized environments, testing, CI-oriented workflows.
- **Soft skills**: effective communication with stakeholders, teamwork in cross-functional environments, structured problem solving.

Experience

Senior Machine Learning Engineer at *Cy4Gate (ELT Group)* **2023.11–present**

◇ Initially onboarded as an external consultant via Frontiere (2023.11–2024.08), then converted to full-time employee (2024.09–present). **Own the end-to-end ML lifecycle** for the RTA SIEM platform in **on-premise** environments: from data ingestion and feature engineering to model training, validation, and **production serving**. Designed and deployed **hybrid batch/online anomaly detection pipelines** using **scikit-learn** and **River**, delivered through a **containerized model serving infrastructure** with **event-driven streaming inference** built on **Kafka** and **Elasticsearch** for observability.

Lecturer (Freelance) at *ITS Lanciano* **2024.04–2024.06**

◇ Teaching **object-oriented programming** in Python.

Senior Machine Learning Engineer (Freelance) at *2oTab* **2023.08–present**

◇ Designed and delivered the *CONNECT* service, owning the full lifecycle of an **NLP-driven semantic search platform** built with **Python**, **Hugging Face** transformer models, and a **Qdrant** vector database. Implemented **vector-based retrieval**, similarity scoring, and **attention-driven explainability** for scientific literature analysis at scale.

◇ Led the development and fine-tuning of **YOLO**-based visual recognition models for the CNR *AUTOMA* project, implementing **data augmentation** and specialized training pipelines to enable **real-time detection of invasive marine species** in Italian waters.

Middle Machine Learning Engineer at *Frontiere* **2023.03–2023.11**

◇ Designed and deployed **containerized ML services** using **Docker**, including *PadelCam*: a **computer vision inference pipeline** running on **Raspberry Pi**, combining **YOLO**-based human detection, automated region-of-interest extraction, and **CNN**-driven image classification for real-time court analysis.

◇ Built an **NLP-based email classification system**, leveraging **BERT** embeddings and a supervised classification pipeline to automatically distinguish *spam* from *non-spam* messages in production environments.

Junior Machine Learning Engineer at *Aesys* **2019.11–2023.03**

◇ For **Generali Assicurazioni**: owned the design and deployment of **NLP and OCR-driven document intelligence pipelines**, combining **SpaCy**, **Tesseract**, **CRNN**, and **Vision/Textract** services with **computer vision** techniques. Delivered **document classification models** using **scikit-learn** and **Keras**, fully deployed in **GCP** with scalable serving architecture.

◇ For **Levis**: developed and optimized **B2C sell-out forecasting models** using **scikit-learn**, with automated **hyperparameter tuning** via **Optuna** on **AWS**, improving predictive accuracy for retail decision-making.

◇ For **Digital Soft**: led the development of **predictive maintenance** and **demand forecasting pipelines** using **scikit-learn** and **Keras**, integrating **model explainability** through **SHAP** to ensure transparent and interpretable results in production.

Research Fellow at *University of L'Aquila* **2019.04–2019.11**

◇ **Deep Reinforcement Learning** applied to Cyber Security through **Mininet** simulations, using **Keras** with a **Q-Learning** algorithm. Malware analysis from `.pcap` files sourced from *Malware Traffic Analysis*.

Education

Master's Degree in Computer Engineering (110/110) *University of L'Aquila*. **2016–2018**

◇ Thesis: *Deep Reinforcement Learning applied to Cyber Security*.

◇ Main study areas: software engineering, algorithms and data structures, advanced databases, machine learning.

Publications & Talks

MTA-KDD'19: A Dataset for Malware Traffic Detection *CEUR-WS* **2020**

◇ Updated dataset for malware traffic analysis, built through collection, cleaning and preprocessing of large volumes of network traffic to train machine learning models.

A Spherical Directional Anemometer Sensor System *MDPI* **2017**

◇ Proposal and analysis of a compact directional anemometer without mechanical moving parts, based on differential pressure measurements using inductive transducers.

Technical articles and blog content available at: lucadivita.it

The full list of my talks is available at: sessionize.com/lucadivit

Hobbies

- *Training and running:* gym workouts and running as part of my wellbeing routine.
- *Reading, studying & blogging:* fiction and technical reading, continuous study, and writing technical articles.
- *Tech community:* active contributor, organizer, and speaker.
- *Gaming:* casual gamer in my free time.