

## DAVIDE MONACO

### OBJECTIVE

Data Scientist & Engineer with an automation background, delivering end-to-end ML in Python across Hadoop/Spark and cloud. Builds robust models, designs intelligent systems, and engineers data pipelines; currently researching VOC detection with genetic algorithms (patent potential). Grounded in statistics, optimization, and systems modeling, with a strong focus on innovation and continuous learning.

### SKILLS & ABILITIES

- ✓ Programming: Python, R, Java, MATLAB
- ✓ Data & ML Frameworks: PyTorch, TensorFlow, Keras, Scikit-learn, Spark, Hadoop, Kafka, PySpark
- ✓ Data Handling: SQL (MySQL), NoSQL (MongoDB), ETL//ELT Pipelines
- ✓ Visualization: Tableau, Matplotlib, Seaborn
- ✓ Tools & Platforms: Git, GitHub, Jupyter, Google Colab
- ✓ AI & Algorithms: CNN, RNN, BERT/GPT, Reinforcement Learning (Q-learning, SARSA), Genetic Algorithms, Symbolic Regression
- ✓ Mathematics & Statistics: Linear Algebra, Optimization, Hypothesis Testing, Bayesian Inference, DAGs, Regularization
- ✓ Other: HTML, CSS, XML, GDPR compliance, Network protocols (TCP/IP, MQTT), Control Systems (PLC, feedback, stability analysis)

### EXPERIENCE

#### **Data Scientist & Research Engineer (Thesis-based Research) NanoTech Analysis Srl (NTA), Turin | Nov 2024 – Present**

- Conducting advanced research on volatile organic compounds (VOCs) using genetic algorithms to develop a predictive model from raw data collected via MEMS/NEMS devices.
- Developed a custom Python framework from scratch to solve identification tasks using evolutionary computing, with patent-pending potential.
- Applied data preprocessing, feature selection, and optimization techniques in a high-stakes R&D environment.

**Skills:** Python - Genetic algorithm - Optimization - Numpy

### TECHNICAL PROJECTS

#### **Deep Learning Models on Resource-Constrained Hardware (Team Project)**

- Implemented Neural Architecture Search (NAS) with training-free metrics to build efficient CNNs for the Visual Wake Word dataset.
- Achieved up to 82% accuracy on ultra-lightweight architectures using mutations and random search.

**Skills:** Python - Deep Learning - ML4IoT - Convolutional NN - Distilled NN

#### **Code Review & Project Workflow Analysis using Git Data (Team Project)**

- Utilized Large Language Models (LLMs) to summarize commits, analyze workflows, and classify developer intent.
- Applied prompt engineering and natural language interpretation for better CI/CD pipeline efficiency.

**Skills:** Large Language Models - Continuous Integration - Prompt Engineering - Git-Hub

### Genetic Programming Toolbox – Symbolic Regression (Solo Project)

- Built a symbolic regression framework using genetic algorithms in Python for mathematical function discovery.
- Focused on extensibility and optimization with dynamic bounding and operator control.

**Skills:** *Genetic Algorithm - Optimization algorithm - Local Search*

### Sentiment Analysis on Twitter Dataset (Team Project)

- Built and evaluated a pipeline of NLP preprocessing and supervised learning algorithms to classify tweet sentiment.

**Skills:** *Supervised Learning - Data Cleaning – Pipeline*

### Student Career Simulation & Animal Ecosystem Simulator (Solo Projects)

- Designed simulators for academic performance modeling and ecological trait-based survival dynamics using Python.
- Modeled decision-making, evolution, and migration dynamics.

**Skills:** *Simulator - Dynamic distribution - Modeling problem*

### Pirelli Innovation Strategy (Team Project)

- Developed a luxury product concept for exclusive Pirelli wheels inspired by motorsport icons.
- Delivered a business model innovation strategy to elevate brand prestige in high-end markets.

**Skills:** *Innovation Management - Business Model*

---

## EDUCATION

### **M.Sc. in Data Science and Engineering**

***Polytechnic University of Turin (Taught entirely in English) (Turin, Italy) –***

**Graduating: Sept 2025**

**Thesis:** Design and development of intelligent algorithms for the analysis of raw data acquired through highly miniaturized devices for VOC identification.

**Coursework:** Machine Learning, Deep Learning, Computational Intelligence, Big Data, Data Engineering, Probability & Statistics.

### **B.Sc. in Automation Engineering**

***Polytechnic University of Bari – Jul 2021***

**Thesis:** Comparative analysis of channel models for Internet of Drones communication.

## CERTIFICATIONS

**University of Michigan:** Programming for Everybody (Python) & Python Data Structures

**Johns Hopkins University:** HTML, CSS, and JavaScript for Web Developers

**AWS Cloud (in progress):** Introduction to IT and AWS Cloud

---

## LANGUAGES

**Italian** – Native

**English** – Fluent (B2)