

# SERGIO VERGA

## AI Engineer | Machine Learning & Data Engineer

✉ sergioverga3005@gmail.com

🔗 SergioV3005

🌐 sergio-verga

📍 Milan, Italy (EU citizen), open to Switzerland

AI / ML Engineer with a strong quantitative background (Physics + Data Science) and hands-on experience across **machine learning**, **computer vision**, **multimodal systems**, and **generative AI**. Experienced in both **research-driven** and **industrial environments**, working on robotics, LLM pipelines, KPI analytics, and data-centric ML workflows. Strong focus on **data quality**, **reproducibility**, **evaluation**, and **clear communication** between research and business stakeholders.

## EXPERIENCE

### External Research Collaborator — AI for Robotics & Industry 5.0

Milan, Italy, Intelligent Sensing Lab (UniMiB)

Python PyTorch Computer Vision Multimodal Data Robotics Evaluation

📅 Mar 2025 – Oct 2025

- Research project: “*Data Collection and Analysis of Human–Machine Interaction in Industry 5.0*” (Cod. 25CE114), part of PRIN 2022 project **HOMEY**, funded by EU NextGenerationEU.
- Collected and analyzed **human–robot interaction datasets** using the **LeRobot** framework across multiple teleoperation interfaces.
- Extended **ACT imitation learning policies** with visual task encodings and multi-task capabilities.
- Benchmarked **wearable-sensor-based control** (Movella Xsens Dot) against leader-arm teleoperation with 1-to-1 kinematic mapping.
- Designed evaluation protocols to compare performance across tasks, modalities, and interaction setups.
- Research output currently under **conference submission**.

### Data Scientist — Generative AI & Analytics

Berlin, Germany, KNIME

Python KNIME LLMs Data Pipelines Explainable AI

📅 Oct 2024 – Dec 2024

- Worked in the **Evangelism Team**, focusing on Generative AI and analytics integration via KNIME workflows.
- Developed a complete **data app** assembling the KPI dashboard featured in KNIME's official blog “*6 CEO KPIs and How to Measure Them*”.
- Built **multimodal local LLM workflows** to automatically interpret KPI reports; presented at **5+ international conferences**.
- Designed **GenAI-powered feature engineering pipelines** and documented them through technical blogs.
- Co-facilitated the **KNIME GenAI Learnathon** at the University of Pisa with **100+ students**.

### Research Intern — Microelectronics

Milan, Italy, Microelectronics Group, University of Milan-Bicocca

MATLAB Signal Processing Modeling Data Analysis

📅 Oct 2022 – Feb 2023

- Conducted research on electronic signal conditioning systems for proton sound detectors.
- Developed and validated numerical models in MATLAB to analyze noise, stability, and frequency response.
- Worked on low-level system concepts, simulation, and structured experimental validation.

## ACHIEVEMENTS

### Winner — Machine Learning Challenge

ML Hackathon, University of Milan-Bicocca

Machine Learning Pipelines Explainability

📅 2023

- Delivered an end-to-end ML solution including data preprocessing, model training, evaluation, and deployment.
- Emphasis on data quality, explainability, and reproducible experimentation.

## SPEAKING & WORKSHOPS

### Speaker — Lean Six Sigma Integration with KNIME

Conference Talk, Accademia Italiana del Sei Sigma (AISS) / KNIME

Public Speaking Process Automation

📅 2025

- Presented a real-world project on integrating **Lean Six Sigma methodologies** with **KNIME-based analytics**.
- Event hosted at Università degli Studi Guglielmo Marconi (Rome).

### Workshop Contributor — KNIME Machine Learning Learnathon

Hands-on Workshop, Fedrigoni Group

Workshops Stakeholder Interaction

📅 2025

- Participated in a full-day, hands-on Learnathon with cross-functional teams (Supply Chain, Finance, HR).
- Supported practical adoption of analytics and ML tools in real business scenarios.

## Speaker — GenAI & AI-driven Analytics

Technical Talk, KNIME Data Connect, Bologna

Technical Communication Analytics

📅 2025

- Co-presented a session on **AI-driven analytics workflows** and applied Generative AI.
- Engaged with a mixed audience of practitioners, consultants, and technical professionals.

## Workshop Co-Facilitator — GenAI Learnathon

Hands-on Workshop, KNIME / University of Pisa

GenAI Teaching

📅 2025

- Co-facilitated a full-day GenAI Learnathon with **106 students**.
- Supported hands-on experimentation with LLM-based analytics workflows.

## Workshop Assistant — Generative AI & Analytics

Corporate Workshop, Gucci, Florence (with Rosaria Silipo)

Industry Training Stakeholder Interaction

📅 2025

- Assisted delivery of an internal workshop on **Generative AI and data analytics**.
- Supported live demos, exercises, and technical discussions with industry stakeholders.

## SELECTED PROJECTS

- **Automatic KPI Interpretation with Multimodal LLMs** — LLM workflows in KNIME for explaining KPI trends and anomalies from dashboard inputs.
- **Vision-Language-Action Models for Robotics** — Multimodal agents for robotic manipulation; benchmarking teleoperation interfaces and extending ACT models.
- **Semantic Segmentation on CamVid** — Full training pipeline comparing custom U-Net-style model with MobileNetV2 and DeepLab.
- **LLM Belief Bias Evaluation** — Evaluation of belief bias in local LLMs using syllogistic reasoning tasks.
- **Emotion Analysis of Stock Tweets** — Multi-class emotion classification and topic modeling with classical and transformer-based NLP models.
- **Hourly Traffic Time-Series Forecasting** — Comparison of SARIMA, UCM, and deep learning models.
- **Text2Topic (Multi-label NLP)** — Large-scale topic classification on ~200k articles with embeddings and LLM baselines.
- **Abstractive Summarization (XSum)** — Benchmarked GRU, Transformer, and LLM approaches with ROUGE and BERTScore.
- **Portfolio Optimization Benchmark** — Classical Markowitz vs. LSTM-based forecasting on historical Nasdaq data.
- **Lex-RAG Lab** — Lexical RAG pipelines (BM25) for document retrieval and QA.
- **Jailbreak Game with LLMs** — Text-based game exploring prompt robustness and instruction-following limits.
- **BSc Thesis — Chopper Preamplifier Modeling** — MATLAB modeling of noise and stability trade-offs in analog electronics.

## EDUCATION

### Master's Degree in Data Science

University of Milan-Bicocca

📅 110/110 with honors

📅 Oct 2023 – Oct 2025

- Thesis: *"Evaluating Teleoperation Interfaces and Multi-Task Learning Approaches for Vision-Based Robotic Manipulation"*.

### Bachelor's Degree in Physics

University of Milan-Bicocca

📅 92/110

📅 Oct 2019 – Feb 2023

- Thesis: *"MATLAB Modelling of a Chopper Preamplifier for Proton Sound Detectors"*.

## PUBLICATIONS & TECHNICAL WRITING

### Automate KPI Report Interpretation and Insights with GenAI & KNIME

Technical Article, KNIME Blog

GenAI LLMs Analytics

📅 2024

- Describes how to operationalize **multimodal local LLMs** to automatically interpret KPI dashboards inside KNIME.
- Focus on explainability, structured outputs, and integration into production analytics workflows.

### 6 CEO KPIs and How to Measure Them

Technical Article, KNIME Blog

KPIs Data Modeling Business Analytics

📅 2024

- Practical guide to defining, computing, and validating key executive KPIs.
- Covers data sourcing, aggregation logic, and reproducible analytics design.

Visualization   GenAI   KNIME

📅 2024

- Walkthrough of KPI visualization workflows enhanced with **GenAI-assisted analytics**.
- Demonstrates practical integration of AI into business-facing dashboards.

LANGUAGES

---

Italian — Native   English — C1   German — A2