

# TOMMASO CRIPPA

## High Performance Computing Engineering Student

✉ crippa.tommaso@gmail.com

🐙 Crippius

🌐 tommaso-crippa

HPC Engineering student with hands-on experience building data pipelines and ML-backed analytics tools. Proficient in Python, SQL, and Kafka, seeking to apply these skills in scalable data infrastructure and ETL processes.

## EDUCATION

### Double Master's Degree in High Performance Computing

Politecnico di Milano - University of Luxembourg

📅 29.25 / 30

📅 2024 - present

- Selected for the exclusive EUMaster4HPC double-degree programme, in which I participated in extracurricular HPC challenges, an international summer school and workshops with industry leaders

### Bachelor degree in Computer Engineering

Politecnico di Milano

📅 110 cum laude / 110

📅 2021 - 2024

- Won "Migliori Matricole" scholarship for being one of Politecnico di Milano's best freshmen students of the 21/22 season.

Relevant courses completed: Streaming Data Analytics, Time Series Analysis, Artificial Neural Networks and Deep Learning

## EXPERIENCES

### Master's Thesis Student @ MEGWARE Computer Vertrieb und Service GmbH

Chemnitz, Germany

🔗 `HPC` `Machine Learning` `Time-Series` `SLURM` `Linux`

📅 February 2026 - present

🐙 Crippius/hpc-bottleneck-detector

- Building a deployed analytics service that ingests ~5,000 data points per minute from 100+ HPC metrics, applies ML-based anomaly detection, and delivers actionable bottleneck reports to HPC users.
- Designing the end-to-end data architecture: from the initial data ingestion and transformation to the automated ml pipeline to enable optimization recommendations without source-code instrumentation.

### Operations & Logistic @ Google Developer Group On Campus - Polimi

Milan, Italy

🔗 `AI` `Cloud` `Mobile`

📅 August 2024 - August 2025

- Organised and executed 5+ technical events over 12 months on AI, Cloud Computing, and Mobile Development, including a hackathon with 150+ participants, doubling community membership over the year.

## PROJECTS

### Benchmarking Detectors in Streaming Machine Learning

🔗 `Python` `capympoa` `pandas` `numpy` `matplotlib` `Colab`

🐙 Streaming-Data-Analytics/Benchmarking-Detectors-in-SML

- Benchmarked concept drift detectors across 4 streaming datasets (~3M data points total) covering abrupt, gradual, incremental, and recurrent drifts, evaluating accuracy, Kappa, and runtime to identify best-fit detectors per scenario.

### Active Learning for Social Media Data Classification

🔗 `Python` `tensorflow` `matplotlib` `numpy` `pandas` `Colab`

🐙 Crippius/al\_project

- Built an active learning pipeline selecting the most informative tweets from 325,000 posts using different selection techniques, reaching F1 ~0.8 with only 1,000 labels available through hyperparameter tuning.

### Whatsapp Wrapped

🔗 `Python` `matplotlib` `pandas` `fpdf` `Flask` `HTML` `CSS` `Javascript` `Vercel` `Render`

🐙 Crippius/whatsapp-wrapped

- Built a full-stack web app that transforms WhatsApp chat exports into personalised analytics reports (message trends, word frequency, response times), deployed on Vercel and Render with 100+ users since launch.

### Early Alzheimer Detection with Deep Learning

🔗 `Python` `pytorch` `keras` `nilearn` `C++` `libtorch` `clinica` `SLURM`

🐙 justvictor/Early\_Alzheimer\_Detection\_DL

- Developed a deep learning model for early Alzheimer's detection, achieving Macro-AUC of 0.9197 on the test set, improving upon the original paper (0.80), with full data preprocessing and augmentation pipelines run on the MeluXina supercomputer.

## SKILLS AND LANGUAGES

Tech Stack: Python (Advanced), Java, Kafka, SQL (Advanced), Spark

Infrastructure: Git (Advanced), DVC, Docker, AWS

Languages: Italian (Native), English (Proficient - IELTS 7.5), German (Beginner - A2)