Vangelis Kourlitis PhD

Data Engineering | Machine Learning | ex-CERN

PROFILE

Data scientist and software developer with 10 years experience in various technical projects and leadership positions at CERN. Specialized in developing scalable pipelines in big data analytics and machine learning domains, with a proven track record of optimizing workflows and delivering impactful solutions. Skilled leader with extensive experience mentoring teams, managing R&D projects and authoring and reviewing technical reports. Now seeking to transition into environmental industry to apply advanced technical expertise to commercial challenges.

CONTACT

Email: kourlitisV@hotmail.com
Tel: +30 6949552002
LinkedIn: vangelis-kourlitis
GitHub: ekourlit

SKILLS

Advanced
Python • C++ • Git

Intermediate
Unix

Novice
Docker • GCP • AWS • SQL • CUDA

EDUCATION

UNIVERSITY OF SHEFFIELD
PhD in Physics
2015 - 2019 | Sheffield, UK

ARISTOTLE UNIVERSITY OF THESSALONIKI
B.Sc. in Physics
2009 - 2015 | Thessaloniki, GR

EXPERIENCE

TECHNICAL UNIVERSITY OF MUNICH & CERN | Geneva, CH
Analysis Model Group Coordinator
Oct 2023 - Dec 2024

- Managed a geographically distributed 120-member group overseeing data analysis software (∼1M LOC) and data formats ensuring project milestones aligned with organizational goals.
- Drove the adoption of a lightweight data format ahead of schedule, offering 3x reduction in data storage costs.
- Rolled out streamlined software configuration paradigm, reducing onboarding time from weeks to hours.
- Planned and mentored multiple fixed price/time projects (€10-20k, 3-6 months) of early-career developers, delivering innovative software products on time and fostering professional growth.

Data Science Researcher
Mar 2023 - Sept 2023

- Awarded €40k grant to lead a 10-member team to modernize legacy C++ data analysis tools with scalable array-based implementations, achieving 4× higher throughput and aligning with industry data science standards.

Python libraries: `Awkward`, `Dask`, `Numba`, `CuPy`

ARGONNE NATIONAL LABORATORY | Chicago, US
Data Science Researcher
Nov 2019 - Feb 2023

- Led a team of 7 researchers in applying advanced analysis algorithms and ML for complex data classification, increasing experimental reach by 12%.
- Developed MLP surrogate with physics-constrained loss terms to approximate deterministic geometrical mathematical functions and offload time-series analysis computations on GPUs.
- Benchmarked novel accelerator SambaNova RDU against NVIDIA A100 GPUs on a variational autoencoder for anomaly detection, achieving a 10× increase in training throughput.
- Facilitated collaboration among 10 international teams by establishing and documenting standardized methodologies for robust AI-driven results.
- Stress-tested Google Cloud Platform's readiness and scalability for large-scale data analysis workflows handling datasets of O(100 TB).

Python libraries: `PyTorch`, `TensorFlow`, `MLflow`, `Ray`, `PySpark`

UNIVERSITY OF SHEFFIELD | Sheffield, UK
Doctoral Researcher
Nov 2015 - Oct 2019

- Authored the first technical report at CERN to publicly release a complete statistical model, setting a precedent for reinterpretation and transparency in the field.
- Developed and maintained C++ data analysis software for high-throughput computations across the full development lifecycle.
- Taught Python programming, scientific computing and ML fundamentals to MSc students through hands-on lab sessions.

Python libraries: `NumPy`, `pandas`, `SciPy`, `scikit-learn`