Francesca Gerardi

Data Scientist, PhD in Astrophysics

16/07/1994 | Citizenship: Italian

Viale Evaristo Stefini 2, 20125, Milano, Italy

fgerardi.uk@gmail.com | frgerardi.github.io | frgerardi | frgerardi | 0000-0002-2743-244X

I started my career as a Data Scientist a couple of years ago, after gaining experience in Bayesian Inference and Machine-Learning during my PhD.

Skills and strengths

Programming	Python: Advanced SQL: Intermediate Fortran: Intermediate C++: Basic R: Basic HTML: Basic
Operating Systems	Ubuntu, Windows, MacOs
Platforms	Databricks, GCP
Deep Learning	Tensorflow, Keras
Version control	git, Github, Bitbucket Google internal tool
Inference	Causal Inference, Simulation-based inference
Bayesian Statistics	Nested Sampling, MCMC, Stan *(pyStan)*
Software	Google Workspace, Microsoft Office
Personal strengths	Self-motivated, Learner, Persistent, Curious
Languages	Italian: Native English: Fluent (CEFR C1) French: Basic

Latest Conferences

[2025] Invited at the 25th International Conference for Robust Statistics, "Providing robust causal analyses for business use cases".

[2022] Speaker at DESI Collaboration Meeting in Berkeley, presented the [2022] paper.

Additional Projects

Cyber-security project on Deepfakes - colab. with NCC Group. Publication Link [first year of PhD, tool Faceswap (github link)]

Doctoral Activities

- Organizer of Journal Club (*2021-22*)
- Organizer of PhDs discussion meetings (*2020-21*)

Personal Interests

Sport • Nature & cultural trips • Gaming • Loud classical and pop music • Art • Cooking and good food

Experience

Data Scientist @ Accenture — *Milan, IT*

Accenture S.p.A. – Strategy & Consulting - Center for Advanced AI — *Jun 2024 - Ongoing*

- **Causal Inference project** – Resources industry:
  I built a causal inference framework which aims to model the cause-effect relationship between customer satisfaction and behaviour, e.g. churn, in order to quantify the economical value of the population-level NPS. A focal point of the project has been the definition of a pipeline to construct a robust causal graph, accounting for both prior knowledge and significance against data.
  Python libraries: causal-learn for causal discovery, DoWhy for GCM modelling
- Tools: Python / Platforms: Databricks (w Optuna, mlflow)

Data Scientist @ Versace — *Milan, IT*

Versace IT - Inventory and Demand Planning Dept. — *Sep 2023 - Jun 2024*

- I built reporting tools for team members, assisting in forecasting and distribution operations.
- I worked on product types' seasonality modelling, with the adoption of clustering techniques.
- Tools: Power BI (DAX), Python, Excel / Platforms: GCP, Oracle OBIEE

Data Science Intern @ Google — *London, UK*

Google UK — *Aug 2022 - Nov 2022*

- I developed extensions to Natural Gradient Boosting (github link), a Machine-Learning algorithm for **probabilistic regression**. These were implemented in production-ready, unit-tested code and improved the performance of the ML model used by my host team by 10%
- Tools: Python (w multiprocessing), version control (internal tools) / Platform: Google Workspace

Teaching Assistant @ London Business School — *London, UK*

London Business School - 'Python for Finance' course — *Apr 2022 - Jun 2022*

Postgraduate Education

PhD in Astrophysics — *London, UK*

UCL, Department of Physics and Astronomy, Cosmoparticle Initiative — *Oct 2019 - Jul 2023*

- Thesis (Link): "Simulation-based inference and data compression applied to cosmological problems"
- **Statistics**: Simulation-based inference, Bayesian hierarchical modeling, Population-level inference
- **Machine-learning**: data compression and probability density estimation
- Tools used: Python (w mpi4py/multiprocessing), Deep Learning libraries (Tensorflow, Keras)
- Access to supercomputer clusters NERSC (U.S. Energy Dept), Hypatia and Splinter (UCL)
- Member of DESI (Dark Energy Spectroscopic Instrument) International Collaboration since *Dec 2021*

Master's Degree in Astronomy: 110/110 cum laude — *Padua, Italy*

Università degli Studi di Padova — *Oct 2016 - Oct 2018*

- Thesis (Link): "Non-parametric reconstruction of cosmological functions", **Erasmus+** in Leiden (NL)

Peer-Reviewed Journal Articles

[2024] Francesca Gerardi, *et al*. Optimal data compression for Lyman-α forest cosmology. MNRAS, Volume 528, Issue 2, Feb 2024 Link.

[2022] Francesca Gerardi, *et al*. Direct cosmological inference from three-dimensional correlations of the Lyman α forest. MNRAS, Volume 518, Issue 2, Jan 2023. Link

[2021] Francesca Gerardi, *et al*. Unbiased likelihood-free inference of the Hubble constant from light standard sirens. Phys. Rev. D, 104:083531, Oct 2021. Link

[2019] Francesca Gerardi, *et al*. Reconstruction of the Dark Energy equation of state from latest data: the impact of theoretical priors. JCAP, 07:042, 2019. Link