Francesca Gerardi

Data Scientist
16/07/1994 | Citizenship: Italian
fgerardi.uk@gmail.com | frgerardi.github.io | frgerardi | frgerardi | 0000-0002-2743-244X

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

Skills and strengths

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

Undergraduate Studies

Bachelor's Degree in Astronomy 105/110
Padua University, Padua, Italy *(Oct 2013 - Sept 2016)*

Doctoral Activities

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

Extracurricular Activities

- Volunteer afterschool with foreign kids (*2009*)
- Private Maths and Physics classes (*2015-2019*)
- Animator in summer camp (*Summers 2007, 2012*)
- Articles writing for school website (*2010-2012*)
- Volleyball (*2009-2014*), one-year team captain

Personal Interests

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

Experience

Data Scientist @ Versace — *Milan, IT*

Versace IT — *4 Sep 2023 - Ongoing*

- Inventory and Demand Planning Dept.
- Assisting team members in forecasting and distribution operations, framing operational problems towards the deployment of data-driven solutions and the definition of more suitable mathematical/statistical analyses.
- Python, SQL • Google Cloud Platform, BigQuery • Power BI, Looker Studio, Excel

Data Science Intern @ Google — *London, UK*

Google UK — *22 Aug - 25 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%
- Software and modules used: Python, version control (Google internal tools), numpy, matplotlib, scikit-learn, pandas, multiprocessing, Google Workspace

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

London Business School — *Apr 2022 - Jun 2022*

- Courses: 'Python for Finance', with exam marking
- Assisted students ($\sim$ 50/class) during lecture coding laboratories
- Software and modules used: Python, Spider, conda, numpy, pandas, seaborn

Postgraduate Education

PhD in Astrophysics [3.5yrs long] — *London, UK*

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

- Thesis work: "Simulation-based inference and data compression applied to cosmological problems"
- Bayesian statistics and data analysis applied to high dimensional problems and Big-Data in the context of Gravitational waves and Large Scale Structure science
- Simulation-based inference, Bayesian hierarchical modeling and Population-level inference
- Machine-learning, *e.g.* for data compression (regression) and probability density estimation
- **Data Intensive Science (CDT DIS) Project on Deepfakes** - colab. with NCC Group. Publication Link
- Software and modules used: Python, conda, Jupyter notebook, numpy, scipy, matplotlib, pandas, seaborn, scikit-learn, pingouin, getdist, mpi4py, multiprocessing, 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 work: "Non-parametric reconstruction of cosmological functions", **Erasmus+** in Leiden (NL)

Peer-Reviewed Journal Articles

[2023] 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