Francesco **Parisio** Ph.D.

Github LinkedIn Email Telephone

EXECUTIVE SUMMARY **Lead Data Scientist** with 5+ years of experience in machine learning and python. Doctoral graduate in computational geomechanics from the Laboratory of Soil Mechanics at EPFL. Leadership record in scientific, engineering, and software development projects. Won multiple highly competitive grants.

Experience

LEAD SCIENTIST | Spanish National Research Council

Sep 2021 – Current | Barcelona, Spain

- Machine Learning: developed physics-based hyper-tuned (Bayesian Optimization) deep learning models (Keras) to forecast energy production of GW-scale CO2-based geothermal systems (error <2.5%).
- Programming: performed 1k large-scale parallel simulations on a HPC cluster for a speed-up of the simulations of ∼3 OOM.
- Innovation: proposed a methodology to store up to 500Mt/yr of CO_2 worldwide at zero leakage risk.

PRINCIPAL INVESTIGATOR | Technische Universität Bergakademie Freiberg

Mar 2019 - Aug 2021 | Freiberg, Germany

- Machine Learning: developed a platform combining GRU-RNN deep learning models (PyTorch) with a large-scale poro-mechanical simulator (C++) for data-driven mechanical simulations with error <3%.
- Innovation: derived a solution to extract ∼50 MW of electric energy from previously non-productive supercritical geothermal wells.
- Team Management: lead an international team that published 8 ISI research papers in 2.5 years.

ASSISTANT RESEARCHER | Helmholtz Centre for Environmental Research

Jan 2017 - Feb 2019 | Leipzig, Germany

- Programming: derived and implemented the first fracture propagation model for the C++ open-source software OpenGeoSys with benchmark error <1%.
- Collaboration: performed research on computational multi-physics of fractured porous media in a team of software developers, including continuous integration (Jenkins), version control (git) and differential testing.

Machine Learning Projects

RECOMMENDATION OF AMAZON PRODUCTS | Collections, Sklearn, Surprise

- Implemented an SVD model to predict the rating of Amazon products that achieved F_1=0.87, a consistent improvement from a baseline collaborative filtering model where F_1=0.84.

TIME SERIES FORECAST OF CO_2 EMISSIONS | Statsmodels

- Implemented an optimized and cross-validated a SARIMAX model that improved the forecasting performance (SMAPE from 22% to 2.8%) and its reliability (likelihood of large errors <2.5%).

HOUSING NUMBER DIGIT RECOGNITION | Sklearn, Keras, TensorFlow

- Devised the architecture of a convolutional neural network to improve the accuracy on the validation dataset from an 87% baseline to a 91% final value.

Volunteer

STATISTICS WITHOUT BORDERS

- Volunteer as a pro-bono data scientist.

Skills

MACHINE LEARNING

• Deep Learning • CNN/RNN • Ensenble Learning • Gradient Boosting • Autoencoder • AutoML • Time Series • Recommendation Systems

PROGRAMMING

• Python • SQL • C++

LIBRARIES

• Numpy • Pandas • Scipy • Matplotlib • Seaborn • Sklearn • Statsmodels • Keras • TensorFlow • Pytorch • Pycaret

TOOLS/PLATFORMS

• Git • Shell • Docker

SOFT SKILLS

• Innovation • Collaboration • Leadership • Initiative

LANGUAGES

• Italian • English • French • German • Spanish

Education

EPFL, SWITZERLAND

Ph.D. Geomechanics

Aug 2012 - Jun 2016

TURIN POLITECNICO, ITALY

M.Sc. + B.Sc. Civil Engineering

Sep 2005 - Dec 2011

Certificates

MIT, USA

Applied Data Science Program

Oct 2021 - Jan 2022 | **Link**

Grants

EU H2020 MSCA

Sep 2021 – Aug 2023 | €160'932.48

DFG-JSPS

Mar 2019 – Aug 2021 | €285'550