Manuel Duarte de Vasconcelos Silva

Contact:

PASS
Av. da Liberdade, N.67B 4B
1250-140 Lisboa
Portugal

Mobile: +351 96 4525553

E-mail: madusilva@gmail.com

Web: https://mdusilva.github.io/
https://github.com/mdusilva
www.linkedin.com/in/manuel-silva-5191b5149

Personal profile

For the past two years I have worked as a Data Scientist/Python Developer at PASS, a Lisbon based Fintech startup (http://www.passp.com/). The focus of my work has been the development of quantitative based investment models, using Machine Learning/Statistics algorithms, as well as the development of backtesting and automatic order execution software.
Previously I worked as a researcher at CENTRA/SIM (Multidisciplinary Centre for Astrophysics https://centra.tecnico.ulisboa.pt/), developing Data Mining/Machine Learning tools for scientific exploitation of large datasets from the Gaia mission of the European Space Agency (ESA). I have also worked with code for PSF reconstruction in Adaptive Optics in the context of the OPTICON project. Prior to this I was a PhD researcher, modelling the spiral arms of our own Galaxy as traced by runaway stars[1].

Positions

2018 – Present	**Data Scientist / Python Developer** PASS (http://www.passp.com/), Portugal
2013 – 2017	**Researcher** CENTRA/SIM (https://centra.tecnico.ulisboa.pt/), Portugal
2012	**Researcher** University of Hertfordshire, UK
2006	**Sales consultant** Portugal Telecom, Portugal

Projects

2014 – 2017

Opticon – "Optical Infrared Co-ordination Network for Astronomy"

- Development and testing of software for PSF reconstruction in Adaptive Optics (GLAO), in the context of the MUSE-GALACSI instrument
- Tasks: coding of Python package; testing (including profiling); implementation of performance improvements (multiprocessing, cache); implementation of object-oriented interface; write the documentation
- **Python 2.7; C; json; FITS 3.0**

1 http://www.newscientist.com/article/mg21729044.000

2013 – 2014	**Gaia: National Participation in the Data Processing and Analysis Consortium (DPAC)** • Development of data mining/machine learning software for estimation of stellar parameters in large databases in the context of the European Space Agency's (ESA) mission Gaia • Tasks: coding of two Python packages; implementation of Bayesian inference algorithms; implementation of distributed computing solutions (execnet package; Hadoop/Spark); testing; write the documentation • **Python 2.7/3.6; json**

Education

2007 – 2011	**PhD in Astrophysics – University of Herfordshire, UK** • Thesis: *Runaway Stars in the Galactic Halo: Their Origin and Kinematics* • Supervisor: Dr Ralf Napiwotzki

2004 – 2006	**Master degree in Statistics – Universidade do Porto, Portugal**
	• Thesis: *Um processo de risco perturbado: aproximações numéricas à probabilidade de ruína*
	• Supervisor: Drª Margarida Brito
	• Topics:
	  ◦ Stochastic processes
	  ◦ Machine learning
	  ◦ Multivariate statistics
	• Final grade: "Muito Bom" (Very Good)

1998 – 2004	**First degree in Astrophysics (Physics/Applied Mathematics) – Universidade do Porto, Portugal** Final grade: 13

Publications

[1] Silva M. D. V. & Napiwotzki R., "High Galactic latitude runaway stars as tracers of the spiral arms", 2013, MNRAS, 431, 502-510 – Astro-ph: http://arxiv.org/abs/1302.0761v1

[2] Napiwotzki R. & Silva M. D. V., "Runaway and hypervelocity stars. The supernova connection", 2012, MEMORIE della Società Astronomica Italiana – Astro-ph: http://arxiv.org/abs/1109.4116

[3] Silva M. D. V. & Napiwotzki R., "Ejection velocities of high Galactic latitude runaway stars", 2011, MNRAS, 411, 2596 – Astro-ph: http://arxiv.org/abs/1010.3651

IT skills

General:

- OS administration/scripting (Linux and Windows)
- Git/Github
- Visual Studio Code
- misc. applications: Tex, Gimp, Photoshop, MS Office (Word, Excel, Access)

Developed packages:

- rpsfpy: software in Python for PSF reconstruction in Adaptive Optics
- Pysysp: Python package for synthetic stellar photometry : https://pypi.python.org/pypi/pysysp/1.0.1
- MASS: Massive MCMC sampler (Bayesian inference): https://github.com/mdusilva/mass
- Mamuto: Python package for distributed computing in clusters: https://github.com/mdusilva/mamuto

Programming:

- Python, R, C# / .NET, MATLAB, Fortran 90
- parallel/distributed computing: Spark, Hadoop, MPI

Databases:

- MySQL, Postgres SQL, SQLite, Python-SQL, PL/SQL

Web:

- Typescript, Jekyll, Apache Web servers (configuration)

Python:

- Data Science: XGBoost, scikit-learn, Pandas, SPAMS (dictionary learning)
- Visualization/dashboards: dash, Plotly, seaborn
- distributed computing: ZeroMQ, pyspark, execnet
- Web: Flask, requests
- Databases: SQLAlchemy
- Bayesian inference (MCMC): pymc

Languages

Fluent: Portuguese, English
Good understanding: French, Spanish
Basic Greek and German (level A1 certificate)

Other interests

I am a qualified Futsal referee. I like to play Football and Futsal, playing chess, Philosophy, travelling. I am also engaged in Science outreach activities directed at the general public.