



João P C Bertoldo

25, Brazilian
(+33) 06 25 51 62 57
ID 0000-0002-9512-772X

in joaopcbertoldo

March, 2021

joaopcbertoldo.github.io

joao.bertoldo@mines-paristech.fr

joaopcbertoldo

EDUCATION

MSc in Artificial Intelligence, Systems, Data (IASD) ^(a) Paris-Dauphine – PSL University <i>Subjects: Machine Learning, Deep Learning, Optimization, Image Analysis.</i>	2019 – 2020 Paris, France
MSc in Executive Engineering ^(b) <i>Minor: Data Science</i> MINES ParisTech - PSL University <i>Subjects: Software Engineering, Databases, Probability, Statistics, Operations Research.</i>	2017 – 2020 Paris, France
Mechatronics Engineer (5-year degree) ^(c) University of São Paulo (USP) <i>Obtained a 1-year scholarship for an international exchange at the University of Valenciennes (France).</i> <i>Teaching assistant (6 mos.): conducted workshops on numerical analysis using MATLAB.</i>	2013 – 2020 São Carlos, Brazil

WORK EXPERIENCE

Research Intern (6 mos.) / Research Engineer (since Mar. 2021) ^(d) MINES ParisTech - PSL University <i>Working on 2D/3D X-ray tomography segmentation with U-Nets [1, 2].</i> <i>Talked at a seminar at MINES ParisTech, the Journée ISS France, and the week-course CVML 2021.</i>	since Sep. 2020 (6 mos.) Évry (Parisian region), France
Research & Development Intern ^(e) Mindsay <i>Evaluated explainability methods [3, 4, 5] on layperson-maintained NLP models (classification task).</i>	2020 (5 mos.) Paris, France
Data Scientist Intern ^(f) Datadog <i>Analyzed and developed algorithms for time series, tag analysis, and a knowledge graph.</i>	2019 (7 mos.) New York, USA
Software Engineer Intern ^(g) Kelda Drilling Controls <i>Modeled and designed a time series annotation tool for supervised machine learning.</i>	2018 – 2019 (8 mos.) Porsgrunn, Norway

PROJECTS

- tomo2seg** ^(d) Tomography segmentation with deep learning (since Sep. 2020) [6].
- pymdr** ^(a) Implementation of “Mining Data Records in Web Pages” [7] in Python (1 mo.) [8].
- Sunlight simulator** ^(b) Head of the project; developed a 3D motion simulator in Python (9 mos.) [9].
- Smart coffee machine** ^(c) Developed a web server in .NET (C#) and an Arduino controller (C++) (4 mos.) [10].

SKILLS & PERSONAL INTERESTS

English (C1): fluent, worked in English speaking companies for 1 year, obtained TOEFL 117/120 (2020).

French (C2): fluent, read subjects in French for 3 years, worked in French organizations for 1 year.

Portuguese: native.

Programming:

Python ^(d-g) (NumPy, SciPy, Matplotlib, Pandas), **MATLAB** ^(c), **Java** ^(b), **JavaScript** (React, Redux) ^(g).

Databases MongoDB ^(g), InfluxDB ^(g), PostgreSQL ^(b).

Deep learning TensorFlow ^(d), Keras ^(d).

Prototyping Streamlit ^(e), Jupyter Lab ^(d-e), Jupyter Notebook ^(f-g).

Others ImageJ ^(d), Avizo ^(d), GitLab ^(f), Git ^(f-g), Bash ^(f-g), Docker ^(g).

Writing and content creation: Google Suite, LaTeX, Markdown.

Personal interests: skydiving, podcasts (especially about politics), has been in 18 countries.

REFERENCES

- [1] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. In: *arXiv:1505.04597 [cs]* (May 2015). arXiv: [1505.04597 \[cs\]](#).
- [2] Özgün Çiçek et al. “3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation”. de. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. 2016. DOI: [10.1007/978-3-319-46723-8_49](#).
- [3] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. ““Why Should I Trust You?” Explaining the Predictions of Any Classifier”. In: (2016). DOI: [10.1145/2939672.2939778](#).
- [4] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. “Anchors: High-Precision Model-Agnostic Explanations”. In: *AAAI*. 2018.
- [5] Scott Lundberg and Su-In Lee. “A Unified Approach to Interpreting Model Predictions”. In: *arXiv:1705.07874 [cs, stat]* (Nov. 2017). arXiv: [1705.07874 \[cs, stat\]](#).
- [6] *tomo2seg GitHub repository*. [github.com/joaopcbertoldo/tomo2seg](#).
- [7] Bing Liu, Robert Grossman, and Yanhong Zhai. “Mining Data Records in Web Pages”. In: *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, New York, USA: ACM Press, 2003, pp. 601–606. DOI: [10.1145/956750.956826](#).
- [8] *Pymdr GitHub repository*. [github.com/joaopcbertoldo/pymdr](#).
- [9] *Sunlight simulator GitHub repository*. [github.com/joaopcbertoldo/minlight](#).
- [10] *Smart coffee machine GitHub repository*. [github.com/joaopcbertoldo/Mkafeina](#).