João P C Bertoldo

25, Brazilian
(+33) 06 25 51 62 57
0000-0002-9512-772X
joaopcbertoldo
joaopcbertoldo
joaopcbertoldo.github.io
joao.bertoldo@mines-paristech.fr

March, 2021

Education

MSc in Artificial Intelligence, Systems, Data (IASD) [a] — 2019 – 2020
Paris-Dauphine – PSL University — Paris, France
Subjects: Machine Learning, Deep Learning, Optimization, Image Analysis.

MSc in Executive Engineering [b] | ***Minor: Data Science*** — 2017 – 2020
MINES ParisTech - PSL University — Paris, France
Subjects: Software Engineering, Databases, Probability, Statistics, Operations Research.

Mechatronics Engineer (5-year degree) [c] — 2013 – 2020
University of São Paulo (USP) — São Carlos, Brazil
Obtained a 1-year scholarship for an international exchange at the University of Valenciennes (France).
Teaching assistant (6 mos.): conducted workshops on numerical analysis using MATLAB.

Work Experience

Research Intern (6 mos.) / Research Engineer (since Mar. 2021) [d] — since Sep. 2020 (6 mos.)
MINES ParisTech - PSL University — Évry (Parisian region), France
Working on 2D/3D X-ray tomography segmentation with U-Nets [1, 2].
Talked at a seminar at MINES ParisTech, the Journée ISS France, and the week-course CVML 2021.

Research & Development Intern [e] — 2020 (5 mos.)
Mindsay — Paris, France
Evaluated explainability methods [3, 4, 5] on layperson-maintained NLP models (classification task).

Data Scientist Intern [f] — 2019 (7 mos.)
Datadog — New York, USA
Analyzed and developed algorithms for time series, tag analysis, and a knowledge graph.

Software Engineer Intern [g] — 2018 – 2019 (8 mos.)
Kelda Drilling Controls — Porsgrunn, Norway
Modeled and designed a time series annotation tool for supervised machine learning.

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), PostgresSQL (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.

João P C Bertoldo

25, Brazilian
(+33) 06 25 51 62 57
0000-0002-9512-772X
joaopcbertoldo.github.io
joao.bertoldo@mines-paristech.fr
joaopcbertoldo
joaopcbertoldo

March, 2021

References

Mines ParisTech - PSL University

Subject: 3D tomography image segmentation with convolutional neural networks.
Address: 60 Boulevard Saint-Michel, 75272 Paris, France.
Period: 2020 – 2021.

Henry Proudhon, CNRS Researcher
Phone: (+33) 01 60 76 30 70
E-mail: henry.proudhon@mines-paristech.fr
Address: Materials Center, 63 Rue Henri Auguste Desbruères, 91000 Corbeil-Essonnes, France.

Etienne Decenciere Ferrandiere, Senior Researcher
Phone:(+33) 01 64 69 48 09
E-mail: etienne.decenciere_ferrandiere@mines-paristech.fr
Address: Center for Mathematical Morphology, 35 rue Saint-Honoré, 77305 Fontainebleau CEDEX, France.

Mindsay

Subject: explainability methods (Lime, SHAP, Anchor) applied to NLP classification models.
Address: 18 Rue du 4 septembre, 75002 Paris, France.
Period: 2020.

Dimitri Lozeve, Data Scientist
Phone: (+33) 06 14 24 68 24
E-mail: dimitri@lozeve.com

Mathieu Seris, R&D Software Engineer
Phone: (+33) 06 37 18 06 63
E-mail: mathieu.seris@gmail.com

Datadog

Subject: time series, tag analysis, Correlations project, knowledge graph project.
Address: 620 8th Ave, New York, NY 10018, USA.
Period: 2019.

Stephen Kappel, Lead Data Scientist
E-mail: stephen@datadoghq.com

Jonathan Lenaghan, Director Of Engineering
E-mail: jonathan.lenaghan@datadoghq.com

Kelda Drilling Controls

Subject: Bivrost project, full stack web development, time series annotations for machine learning.
Address: Hydrovegen 6, 3912 Porsgrunn, Norway.
Period: 2018 – 2019.

Geir Arne Evjen, Leading Expert at Kelda
Phone: (+47) 906 01 558
E-mail: gevj@kelda.no

Glenn-Ole Kaasa, CEO at Kelda
Phone: (+47) 913 04 775
E-mail: gok@kelda.no

CMM
Centre de Morphologie Mathématique

13 January 2021

Letter of reference for João P C Bertoldo

To whom it may concern,

I met M. Bertoldo in January 2020, when he enrolled in the course "deep learning for image analysis" I teach in the master's program "Artificial Intelligence, Systems, Data," a highly selective Parisian master's degree. He obtained a final score of 33.5/40, ranking among the best students.

From September 2020 until now, I have been co-supervising his final year project, which should end in March. João is working on the automatic segmentation of 3D tomography images. He has already obtained excellent results from the application point of view. In the process, he has suggested several original ideas, and implemented them efficiently – he is proficient in Python. His deep learning skills are already very good, and he is also open to other disciplines. He plans, for instance, to combine deep learning with mathematical morphology to improve his current results. He has a researcher's attitude: he aims at fully understanding what is going on, remaining critical about his work.

Last but not least, João is also motivated and enthusiastic, and his communication skills are excellent. He has a wonderful rapport with people.

In conclusion, I highly recommend M. João P C Bertoldo for a Ph.D. position in image analysis or deep learning.

Etienne Decencière
Directeur de recherche / Research Director

MINES Paristech – Centre de Morphologie Mathématique
35, rue Saint-Honoré – 77305 Fontainebleau Cedex – France
T. (+33) 1 64 69 47 06 – F : (+33) 1 64 69 47 07
www.mines-paristech.fr – www.mines-paristech.eu

Henry Proudhon
CNRS research director
Centre des Matériaux, MINES ParisTech
63-65 rue Henri desbruères, BP 87
F-91003 EVRY Cedex, France
henry.proudhon@mines-paristech.fr

January 18, 2021

Recommendation of Mr. João P. C. Bertoldo

To whom it may concern:

I'm writing this letter to support the application of Mr. Bertoldo. I am a CNRS researcher in Centre des Matériaux MINES Paristech developing new multimodal X-ray-based characterizations for structural materials. I'm leading the BIGMECA chair, a research initiative between MINES ParisTech and Safran to develop AI-based methods in Mechanics of Materials, notably using X-ray images of microstructure and defects. It is in this framework that I proposed Mr. Bertoldo to work on *deep learning for automated segmentation of tomographic images*, and he started in late September 2020.

João was very dedicated to his work and achieved publication-ready results in just three months. During this time, he proved to already be an autonomous thinker and managed his project on his own, asking for guidance whenever needed. I also wanted to mention the excellent programming skills of Mr. Bertoldo; he seems to be able to take down any problem thrown at him. João rank within the top 3% of the students I had to supervise; if I had the funding to propose a Ph.D. thesis to João in his field, it would do it immediately.

João has been a great asset to my group; I did not hesitate to propose to João to contribute to the doctoral course *Computer vision and machine learning for the material scientist*[1], and he accepted enthusiastically.

I very strongly recommend the application of Mr. Bertoldo for a Ph.D. position in data science. Do not hesitate to contact me for further information.

Sincerely,
Henry Proudhon

[1] `http://dms.mat.mines-paristech.fr/Donnees/data05/503-CVML_new_2021.pdf`

September 6, 2019

Re: Letter of Recommendation for João Paulo Casagrande Bertoldo

To whom it may concern:

I was João's manager when he worked as a Data Science Intern at Datadog in New York City from February 2, 2019 through September 6, 2019.

At Datadog, most of João's time was spent helping our team build a new feature that we announced in July and started rolling out to customers in August. Therefore, in addition to exploring data and developing data-based algorithms to support the feature, João also contributed by writing a significant amount of production code and troubleshooting/optimizing issues in our production systems.

João brings a healthy curiosity and restlessness to the problems he works on. He consistently raises ideas for incremental improvements and alternative approaches to the solutions he is implementing. When working with more senior members of our team, he always interrogates their proposals, making sure he has a strong understanding of the rationale for decisions made.

João is comfortable collaborating with whomever he needs to in order to get his job done. He reaches out to engineers both on the same team and on other teams, as appropriate. João is happy to share/present his work with the team to gather ideas and make sure everyone is on the same page.

I was very impressed with João's commitment to personal development. He voluntarily joined with a teammate in taking a course on our learning platform, and he consistently asked for feedback from myself. Most importantly, João listens carefully to feedback and seems intent on acting on it to grow. Needless to say, this is a very valuable trait for any engineer, especially at the beginning of his career.

Sincerely,

Stephen Kappel
Team Lead, Data Science
Datadog