


Jonathan Crabbé

Cambridge, United Kingdom 

07547340408 

jonathan.cr1302@gmail.com 

linkedin.com/in/jonathancrabbe 

jonathancrabbe.github.io 

Born Feb. 1996, Belgium 

I am looking for a research position where I can put my curiosity, knowledge and passion at the service of pushing the limits of AI. We are gradually entering in a phase where humans will increasingly interact with AIs, hence forming human-AI teams. I see an immense potential in these teams to approach cutting-edge scientific and medical problems. My research focuses on making these teams more efficient through interpretability, efficient representation learning and data-centric insights. My diverse educational background (engineering & theoretical physics) gives me a unique angle to tackle these challenges.

Education

2020-

PhD in Applied Mathematics

University of Cambridge

Synopsis. My thesis has a strong focus on developing and improving machine learning approaches to tackle high-stakes domains, such as medicine. My research involves several subfields of AI research, including: *interpretability*, *representation learning* and *data-centric AI*.

Supervisor. Mihaela van der Schaar

Publications. During my 2 first year of PhD, I have published 8 papers introducing SOTA machine learning methods in NeurIPS and ICML.

2018-2019

MASt in Applied Mathematics and Theoretical Physics

University of Cambridge

Description. This master is one of the most challenging courses of math and physics in the world. Succeeding required me to master complex mathematical paradigms such as *differential geometry*, *group theory* as well as *functional analysis*. This formal knowledge perfectly complements the practical skills that I have acquired during my bachelor.

Grade: Distinction (highest grade)

Relevant courses. Black Holes: 89%, Symmetries, Fields & Particles: 88%, Advanced Quantum Field Theory: 87%

2017-2018

M1 in Physics

Ecole Normale Supérieure Paris

Description. This master is a research-oriented course in physics. To achieve an excellent grade in this course, I had to quickly *overcome the challenging transition* between engineering and physics.

Grade. Mention Bien

2014-2017

Bachelor's in engineering **Université Libre de Bruxelles**

Description. This bachelor's in engineering taught me to solve practical problems with a *principled and pragmatic approach*. I also developed a real interest for using *computer science to solve real-world problems*. In addition, I discovered a passion for *teaching* as a teaching assistant.

Grade. The Highest distinction, *first of my promotion* (among 450 students in first year)

Relevant courses. Analysis 1&2: 100% for both, Complements of mathematics and numerical analysis: 95%, Physics: 100%, Oriented Object Programming: 90%, Numerical Analysis: 95%, Probability and Statistics: 95%, Signals and Systems: 95%, Quantum Mechanics: 90%

Experience

JUNE 2022 – SEP 2022

Quantitative Research Intern **G-Research**

Description. 10 weeks of internship here I applied SOTA machine learning approaches to financial data. I have learned to build an entire modelling pipeline to transform raw noisy data into predictive features that can be used to train a machine learning model.

Transferable skills. data cleaning and labelling, feature engineering, modelling of financial markets

2019 – 2020

Research Assistant **Université Libre de Bruxelles**

Description. 1 year of full-time research in black-holes physics and quantum field theory. Working on such an abstract subject developed my ability to think about problems in a conceptual way. I also realised that this conceptual thinking can serve more concrete fields to have a real-world impact.

Transferable skills. autonomy, perseverance, desire to have a real-world impact

2016 – 2020

Content Creator **Clipedia**

Description. Several years of involvement in the teaching project Clipedia. I have created pedagogical videos on YouTube to teach math, physics, and chemistry to young students. The YouTube channel has almost 100k subscribers and this number is quickly increasing.

Please visit clipedia.be for more information.

Transferable skills. communication, pedagogy, ability to adapt a speech to a wide audience

FEB 2020 – MAY 2020

Research Intern **University of Cambridge**

Description. Part-time research internship in the van der Schaar Lab. This was my first exposure to concrete machine learning research. I managed to develop a new interpretability method to translate black-box machine learning models into closed-form symbolic expressions. This led to a paper published in NeurIPS 2020 and a PhD offer that I accepted.

Transferable skills. identify possible improvements in existing methods, implement those improvements, write a scientific paper, articulate a reasoning to convince other researchers

FEB 2018 – JUN 2018

Research Intern Imperial College London

Description. 4 months of full-time research in Modified Theories of Gravitation. This project was a good mix between theoretical and experimental work. I had to study pathological behaviours of gravity theories and assess my findings through numerical simulations. During this first experience as a researcher, I have also learned the exploration-exploitation trade-off familiar to all researchers.

Transferable skills. formulate a problem, design numerical simulations to verify theoretical claims, manage the time on a project with a clear deadline

Skills

- **Coding.** Strong knowledge of the Python Machine Learning toolkit (including Pytorch, Tensorflow, Scikit-Learn, Pandas, Jupyter and Numpy).
 - **Modelling.** Trained machine learning models on various for a wide range of tasks (including classification, regression, clustering and denoising) over my PhD and experience in finance.
 - **Reasoning.** Learned to articulate a sound reasoning through probability theory, statistics, measure theory, functional analysis, group theory, linear algebra and differential geometry.
 - **Proving.** Most of my papers have a strong theoretical components with rigorous original proofs.
 - **Creating.** My unique educational background (engineering, applied mathematics and theoretical physics) allows me to look at a problem through many different angles.
 - **Presenting.** Through many talks and a strong teaching experience, I have learned to present ideas in an inspiring way for a wide variety of audiences.
 - **Working in a team.** Half of my papers are the result of a collaboration with several co-authors.
 - **Autonomy.** The other half of my papers are the result of autonomous work.
 - **Supervising.** I have personally supervised 2 research internships (1 Cambridge MPhil student and 1 Imperial MSc student) and helped 3 PhD students to define their research agenda.
 - **Managing time.** I believe that my time management skills are well reflected by my publication record (8 papers published in top machine learning conferences over the 2 first year of my PhD).
 - **Languages.** French (native), English (IELTS Band 8).
-

Research interests

- AI for Science & Medicine
- Interpretability
- Representation Learning
- Data-Centric AI
- Robust Machine Learning
- Time Series Modelling

I am extremely interested in improving the information flow between machine learning models and humans. Making this flow as natural and precise as possible has 2 immense benefits:

1. Identify weaknesses of ML models (which is not obvious from out-of-sample testing alone)
2. Extract human knowledge from ML models (which is essential in scientific applications)

Most of my papers cover the above research interests with the idea of improving this flow of information between complex ML models and human users.

Publications

- **NeurIPS 2022.** Crabbé, J., van der Schaar, M. (2022). Concept Activation Regions: A Generalized Framework For Concept-Based Explanations.
- **NeurIPS 2022.** Seedat, N., Crabbé, J., Bica, I., van der Schaar, M. (2022). Data-IQ: Characterizing subgroups with heterogeneous outcomes in tabular data.
- **NeurIPS 2022 Benchmark.** Crabbé, J., Curth, A., Bica, I., van der Schaar, M. (2022). Benchmarking Heterogeneous Treatment Effect Models through the Lens of Interpretability.
- **ICML 2022 Spotlight.** Crabbé, J., van der Schaar, M. (2022). Label-Free Explainability for Unsupervised Models. In Proceedings of the 39th International Conference on Machine Learning, PMLR 162:4391-4420.
- **ICML 2022 Spotlight.** Seedat, N., Crabbé, J., van der Schaar, M. (2022). Data-SUITE: Data-centric identification of in-distribution incongruous examples. In Proceedings of the 39th International Conference on Machine Learning, PMLR 162:19467-19496.
- **ICML 2022 Workshop.** Sun, H., van Breugel, B., Crabbé, J., Seedat, N., & van der Schaar, M. (2022). DAUX: a Density-based Approach for Uncertainty eXplanations. arXiv preprint arXiv:2207.05161.
- **NeurIPS 2021 Spotlight.** Crabbé, J., Qian, Z., Imrie, F., & van der Schaar, M. (2021). Explaining Latent Representations with a Corpus of Examples. Advances in Neural Information Processing Systems.
- **ICML 2021 Spotlight.** Crabbé, J., van der Schaar, M. (2021). Explaining Time Series Predictions with Dynamic Masks. In Proceedings of the 38th International Conference on Machine Learning, PMLR 139:2166-2177.
- **NeurIPS 2020 Poster.** Crabbé, J., Zame, W. R., Zhang, Y., & van der Schaar, M. (2020). Learning outside the black-box: the pursuit of interpretable models. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, & H. Lin (Eds.), Advances in Neural Information Processing Systems (pp. 17838--17849). Curran Associates, Inc.

Conferences

- **NeurIPS 2022.** 2 papers accepted in the main conference and 1 benchmark paper
- **ICML 2022.** 2 papers accepted in the main conference and 1 workshop paper.
- **NeurIPS 2021.** Spotlight paper (top 3% paper) on explaining latent representations.
- **Stanford MedAI Series.** Long presentation on Explainable AI for time series.
- **OxML 2021.** Full summer school attendance.
- **ICML 2021.** Spotlight paper on explaining time series models.
- **NeurIPS 2020.** Poster paper on translating black-boxes into closed-form expression.

Awards

- **AVIVA PhD Fellowship** to fund my PhD (2020).
- **ULB Research Assistant Fellowship** to fund my year as a research assistant (2019).
- **Wolfson College Jennings Price** awarded based on outstanding results for my MAST (2019).
- **Labex-ICFP Scholarship** to fund my M1, awarded based on academic excellence (2018).

Teaching

- **Research Supervisor.** Supervision of one Cambridge MPhil thesis in machine learning and one research internship in theoretical physics (2 × 4 months).
- **Physics Teaching Assistant.** Responsible for 1st year pharma physics exercise sessions (1 year).
- **Pedagogy project Clipedia.** Writing, recording and editing videos to teach young students mathematics (4 years).
- **Engineering Teaching Assistant.** Responsible for weekly Q&A sessions for 1st and 2nd year engineering students (1 year). Subjects: Linear Algebra, Geometry, Analysis, Probability and Statistics, General Physics, Quantum Mechanics, Classical Mechanics.