Jules Samaran

google scholar | github

Education

Institut Pasteur Nov. 2021 – Now
PhD student in Machine Learning and Computational Biology *Paris, France*
Devising methods for dimensionality reduction of unpaired multimodal data at the single cell resolution under the supervision of Laura Cantini and Gabriel Peyré.

University of Paris-Saclay Sept. 2020 – May 2021
MSc Mathematics Vision and Learning (MVA) *Paris, France*
Coursework: *Computational Statistics*, *Convex Optimization*, *Probabilistic Graphical Models*, *Kernel Methods for Machine Learning*, *Theory of Random Matrices*, *Biostatistics*

Mines Paristech, PSL Research University Sept. 2017 – May 2021
MSc in Applied Mathematics and engineering degree *Paris, France*

Lycée Pierre de Fermat Sept. 2015 – July 2017
Undergraduate preparatory program *Toulouse, France*

- A post-secondary two-year program in advanced math and physics leading to nationwide entrance examinations to the Grandes Ecoles for scientific studies
- Relevant coursework includes *Linear Algebra*, *Differential Calculus*, *Probabilities*, *Real Analysis* and *Topology*

Internship experience

Research Intern Apr. 2021 – Aug. 2021
Criteo AI Lab *Paris, France*
Research internship in the Recommender Systems Team

- Investigated the use of a user history-conditioned generative model for recommendation and product innovation. *Our paper was accepted at the RecSys Workshop on Recommender Systems in Fashion.*

Visiting researcher Apr. 2020 – Sep. 2020
Osaka University, Institute for Datability Science *Osaka, Japan*
Research internship in the Vision and Language team

- Studied the addition of supervision to the attention mechanism in a Transformer-based Vision and Language pre-trained model. *Our paper was accepted at the ACL-IJCNLP SRW.*

Visiting researcher Oct. 2019 – Mar. 2020
*A*STAR, Institute for Infocomm Research* *Singapore*
Research internship on two projects under the supervision of Foo Chuan Sheng

- Anomaly detection on multidimensional time series. *Our paper was accepted in TNNLS*
- Approximation of gaussian processes with infinitely wide neural networks in a co-training framework.

Data scientist intern in the Experiments team July 2019 – Oct. 2019
Booking.com *Amsterdam, Netherlands*

- Reviewed A/B experiments run by data scientists to enforce statistical guidelines
- Analyzed the application of causal inference algorithms (Anchor Regression and Invariant Causal Prediction) on real-world data collected from both interventional and observational settings

Visiting student researcher in the EECS department Aug. 2018 – Feb. 2019
UC Berkeley *Berkeley, CA*

- Worked on scVI under the supervision of Romain Lopez and Nir Yosef
- Co-authored a paper that was awarded first prize at the ICML workshop on Computational biology

Publications

1. Geert-Jan Huizing*, Jules Samaran*, Anna Audit, Daniele Capocefalo, Gabriel Peyré, Laura Cantini. STORIES: learning cell fate landscapes from spatial transcriptomics. *Accepted in Nature Methods.*

2. Jérémie Kalfon, Jules Samaran, Gabriel Peyré, and Laura Cantini. scPRINT: pre-training on 50 million cells allows robust gene network predictions. *Nature Communications*, 16(1):3607

3. Jules Samaran, Gabriel Peyré, and Laura Cantini. scConfluence: single-cell diagonal integration with regularized inverse optimal transport on weakly connected features. *Nature Communications*, 15(1):7762

4. Jules Samaran, Ugo Tanielian, Romain Beaumont, and Flavian Vasile. What users want? warhol: A generative model for recommendation. In Nima Dokoohaki, Shatha Jaradat, Humberto Jesús Corona Pampín, and Reza Shirvany, editors, *Recommender Systems in Fashion and Retail*, pages 17–30, Cham, 2022. Springer International Publishing

5. Adam Gayoso, Romain Lopez, Galen Xing, Pierre Boyeau, Valeh Valiollah Pour Amiri, Justin Hong, Katherine Wu, Michael Jayasuriya, Edouard Mehlman, Maxime Langevin, et al. A python library for probabilistic analysis of single-cell omics data. *Nature biotechnology*, 40(2):163–166, 2022

6. Astha Garg, Wenyu Zhang, Jules Samaran, Ramasamy Savitha, and Chuan-Sheng Foo. An evaluation of anomaly detection and diagnosis in multivariate time series. *IEEE Transactions on Neural Networks and Learning Systems*, 33(6):2508–2517, 2021

7. Jules Samaran, Noa Garcia, Mayu Otani, Chenhui Chu, and Yuta Nakashima. Attending self-attention: A case study of visually grounded supervision in vision-and-language transformers. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: Student Research Workshop*, pages 81–86, 2021

8. Romain Lopez, Achille Nazaret, Maxime Langevin, Jules Samaran, Jeffrey Regier, Michael I Jordan, and Nir Yosef. A joint model of unpaired data from scrna-seq and spatial transcriptomics for imputing missing gene expression measurements. *arXiv preprint arXiv:1905.02269*, 2019

Talks

- GDR Iasis: Optimal Transport for Machine Learning ..... ENS Lyon, February 2025
- MILES team seminar ..... PariSanté campus, February 2025
- AI in Biology and Health symposium ..... Pasteur Institute, October 2024
- ECCB 2024 (poster) ..... Turku, September 2024
- LEGO-BIM: Machine learning for genomics ..... Lille, November 2023
- Camudi workshop ..... Marseille, November 2023
- Pasteur Institute's Developmental Biology department retreat ..... Annecy, October 2023
- AI in Biology and Health symposium ..... Pasteur Institute, July 2023
- Journées Boris Ephrussi ..... Paris, May 2023
- Workshop DDisc ..... Toulouse, October 2022
- Multimodal data integration course ..... Curie Institute, Paris, September 2022
- ECCB 2022 (poster) ..... Sitges, September 2022

Teaching

Introduction to Artificial Intelligence Spring 2025
Oteria Cyber School, Bachelor 1 *Paris, France*
12-hour course on the fundamentals of AI

Preparation for competitive oral exams ("colles") Spring 2022
Lycée Fénelon, Mathematics (MPSI) *Paris, France*
2 hours per week

Side projects

Owkin & Servier AI 48h Hackathon for Glioblastoma Research February 2025

ARIEL | *Python* February 2019 – June 2019
Project conducted in free-time with Timothée Launay: code available at ARIEL

- Devised and implemented an algorithm that takes as input an image, and tries to draw it as a human would: by drawing lines sequentially on a blank canvas

Facebook Paris 24h Hackathon | *Python* May 2019

- Coded the backend structure of a website providing analysis of interactions within messenger groups

Experimental research project | *Python* Sept 2016 – March 2017

- Solved mathematically and computationally a generalized version of the Wythoff game

Skills

Programming: Python, Java, R, Git, Spark (Pyspark)
Languages: French (native), English (TOEFL: 112/120), Spanish (intermediate), Japanese (beginner)

Extracurricular

Sports

- Rugby: SBLR, Singapore Cricket Club, XV de la Pioche, LOCH
- Judo: black belt (1st dan), national-level competitor while in high school
- Boxing: CalBoxing team in Fall 2018

Volunteer

- Former member of a non-profit association "Col Vert"- Teaching and mentoring underprivileged high school students

Student associative life

- Former member of the Mines Paristech Arts Club - Opera section