

Raphael Poulain

PHD STUDENT IN COMPUTER SCIENCE

439 Smith Hall, Newark, DE

☎ 816-616-6698 | ✉ rpoulain@udel.edu | 🏠 rpoulain.github.io | 🎓 Raphael Poulain

Research Interests

Applied Machine Learning, Artificial Intelligence, Health Data Science, Deep Learning, Electronic Health Records

Education

University of Delaware

PH.D. IN COMPUTER SCIENCE

Advisor: Rahmatollah Beheshti

Newark, DE

2020 - Present

EFREI Paris

M.S. IN ENGINEERING

Villejuif, France

2018 - 2020

EFREI Paris

B.S. IN ENGINEERING

Villejuif, France

2015 - 2020

Work Experience

Pinterest

MACHINE LEARNING RESEARCH INTERN

Remote, USA

May. 2023 - Aug. 2023

- Built an end-to-end pipeline to quantify and analyze demographic biases in text-to-image generative models.
- Curated a list of queries and proposed a set of relevant metrics to quantify fairness and biases.
- Investigated and proposed multiple solutions to mitigate biases issues such as Textual Inversion, Prompt Engineering, and fair fine-tuning through LoRA.
- Improved the performances of the generative model by up to 25% in all bias and fairness metrics with minimal impact on the quality and relevance of the generated images.

Euronext

SOFTWARE ENGINEER INTERN

Paris, France

Mar. 2020 - Aug. 2020

- Built an automated API testing tool using Python and Cucumber during the development of the company's new on-demand market data platform.
- Participated actively in the choice of the technologies and tools used throughout the project.
- Organized the deployment of the tool to serve as the company's main testing engine for newly developed APIs.

Euronext

SOFTWARE ENGINEER INTERN

Paris, France

Apr. 2019 - Sep. 2019

- Built a cartography tool of the Optiq Trading system that allow engineers to better visualize the architecture of the system.
- Designed the Graph Database Model from the choice of the technology to the model itself.
- Realized a WebApp using JavaScript to keep the database up-to-date automatically and to help visualize each connection between Optiq's components.

Research Experience

healthy lAlfe Lab, University of Delaware

RESEARCH ASSISTANT

Newark, DE

Sep. 2020 - Present

- Investigated a novel architecture to extract the graphical structure of medical visits in a transformer-based model.
- Studied a comprehensive Federated Learning approach to promote fairness in the model's prediction on underrepresented populations.
- Investigated a semi-supervised transformers architecture leveraging both labeled and unlabeled patients for few-shot learning on Electronic Health Records.
- Utilized Transformers and Electronic Health Records to perform a multi-target regression for primordial prevention of cardiovascular disease.
- Leveraged a BERT-based architecture to extract the bidirectional representation of patients' medical data by defining medical codes as words, visits as sentences, and the medical history as a document.
- Implementation of a single RNN and GAN-based model to predict obesity status at different time-points in the future.
- Participated in the BARDA Pediatric COVID-19 Data Challenge.

Teaching Experience

CISC106, University of Delaware

TEACHING ASSISTANT

Newark, DE

Fall 20 - Spring 23

Publications

PEER REVIEWED CONFERENCE PAPERS

- [1] **R. Poulain**, R. Beheshti. "Graph Transformers on EHRs: Better Representation Improves Downstream Performance". The Twelfth International Conference on Learning Representations (ICLR), 2024. [Link](#)
- [2] **R. Poulain**, M. F. Bin Tarek, R. Beheshti. "Improving Fairness in AI Models on Electronic Health Records: the case for Federated Learning Methods". 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT), 2023. [Link](#)
- [3] M. Gupta, B. Gallamoza, N. Cutrona, P. Dhakal, **R. Poulain**, R. Beheshti. "An Extensive Data Processing Pipeline for MIMIC-IV". Proceedings of the 2nd Machine Learning for Health symposium (ML4H), 2022. [Link](#)
- [4] **R. Poulain**, M. Gupta, and R. Beheshti. "Few-Shot Learning with Semi-Supervised Transformers for Electronic Health Records". Proceedings of the 7th Machine Learning for Healthcare Conference (MLHC), 2022. [Link](#)
- [5] M. Gupta, **R. Poulain**, TL. T. Phan, H. T. Bunnell and R. Beheshti. "Flexible-window Predictions on Electronic Health Records". Proceedings of the Thirty-Fourth Annual Conference on Innovative Applications of Artificial Intelligence (IAAI), 2022. [Link](#)
- [6] **R. Poulain**, M. Gupta, R. Foraker, and R. Beheshti. "Transformer-based Multi-target Regression on Electronic Health Records for Primordial Prevention of Cardiovascular Disease". 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2021. [Link](#)

Services & Activities

Conference Volunteer IEEE International Conference on Bioinformatics and Biomedicine 2021

Projects

Portfolio Optimization

2020

Portfolio creation following the Markowitz's Optimal Portfolio using the returns predicted by an LSTM-based network.

NHL Players' salaries Prediction

2018

Developed a Random Forest Regression model to predict NHL Players' salaries given their in-game statistics and personal information.

Self Driving Cars

2018

Programmed a parking lot simulation populated by autonomous cars to teach them how to park using a genetic algorithm.

Relevant Coursework

University of Delaware	Artificial Intelligence, Advanced AI, Bioinformatics, Introduction to Data Mining, Advanced Algorithms, Game Theory, Theory of Computation, Introduction to NLP
EFREI Paris	Numerical Analysis Applied to Finance, Financial Risk, Econometrics, Big Data for Finance, Advanced Databases, Deep Learning and Applications

Skills and Certifications

Programming	Python, R, Keras, PyTorch, Tensorflow, SQL, C / C++, Java
Machine Learning	Neural Networks, RNN, Transformers, CNN, GAN
Certifications	Coursera Deep Learning Specialization