

PAUL NOVELLO

AI Research Scientist

@ paul.novello@outlook.fr

+33 628343558

https://paulnovello.github.io/

in paul-novello-a036b1a1

EXPERIENCE

AI Research Scientist

IRT Saint Exupéry

March 2022 – Now

Toulouse, France

- Coordination (as scientific leader) of the creation of a research program on AI for Simulation in Artificial and Natural Intelligence Toulouse Institute (ANITI), gathering both academics (IRIT, Supaero, IMT) and industrials (Airbus, Liebherr, Schaeffler).
- Core member of DEEL (DEpendable and EXplainable Learning) project, part of ANITI. Research on Trustworthy AI (Explainable AI - XAI, Safe AI, Robustness, Uncertainty Quantification), in collaboration with Academics and industrials.
- Lead of the Out-Of-Distribution (OOD) detection research project. Close collaboration with MILA, Airbus, and Renault. OOD for Image classification, Object detection and LLMs
- Lead developer of **odeel**, an open-source library for post-hoc OOD detection on pre-trained Pytorch and Tensorflow deep neural networks.
- 3D prediction of pollutant diffusion in urban areas: Used Fourier Neural Operators (FNO) as Surrogate models for the prediction of pollutant diffusion in 3D meshes of urban areas. Joint work with The French Alternative Energies and Atomic Energy Commission (CEA).
- Supervised a group of 3 PhD candidates during the **CEMRACS** hackathon.
- Oral Communications (NeurIPS Paris, Open Source Experience, UQSay, GDR Mascotnum, Algoritmy, Meetups, Internal...).

PhD Candidate

Ecole Polytechnique, French Alternative Energies and Atomic Energy Commission (CEA) CESTA, INRIA Saclay

Nov 2018 – March 2022

Bordeaux - Palaiseau, France

- Deep Learning for Scientific Machine Learning (SciML) and High Performance Computing (HPC) applications, specifically Computational Fluid Dynamics.
- Large Hyperparameter Optimization on CEA's cluster with 400 GPUs using Slurm.
- Embedding of a Multi-Layer Perceptron (MLP) trained in TensorFlow in a real-world Fortran/C++ simulation code using Tensorflow C++ API. Modification of the original Fortran C++ code to leverage neural network's vectorization.

Data Scientist - Intern then Remote as a Freelance

Intento>

June 2017 – Nov 2018

Bordeaux, France

- Contribution to a server built with Flask, automatically processing MariaDB visitor databases from client websites.
- Implementation of data analysis algorithms: website visitor's fidelity scoring, visitor's profile clustering, and classification based on their navigation path, using TraMineR a Genomics sequence clustering package in R.

EDUCATION

PhD, Applied Mathematics

Ecole Polytechnique

2018 – 2022

Palaiseau, France

- Combining supervised deep learning and scientific computing: some contributions and application to computational fluid dynamics.

MSc, Statistics

Imperial College London

2017 – 2018

London, UK

- Specialization in Machine Learning and Computational Statistics.

Diplôme d'Ingénieur

Telecom ParisTech

2014 – 2018

Paris, France

- Major in Statistical Modeling and Scientific Computing, Signal Processing

SKILLS

Python TensorFlow PyTorch Jax
Slurm Ray Git C++ Fortran
Docker SQL R Matlab

MISCELLANEOUS

- Sports: climbing, hiking, tennis
- Indie rock, contemporary jazz, and techno music. I'm working towards confidently participating in jam sessions with my bass.
- Reading and watching about non-mathematical sciences, economy, geopolitics...
- My first baby girl, Léonie, was born in January 2024!

LANGUAGES

- French: Native
- English: Fluent
- Spanish: Intermediate

PUBLICATIONS

Improving Out-of-Distribution Detection by Combining Post-hoc Detectors

[preprint](#)

📅 2024

- Combination of existing post-hoc Out-Of-Distribution detection scores using multivariate statistical modeling to improve OOD detection performances.

Instance-dependent Approximation Guarantees for Lipschitz Approximators, Application to Scientific Machine Learning.

[preprint - submitted to TMLR](#)

📅 2024

- Deriving exact generalization bounds for Lipschitz approximators based on learning points and train models using certified Deterministic Optimistic Optimization.

Exploring the Link Between Out-of-Distribution Detection and Conformal Prediction with Illustrations of Its Benefits

[preprint - submitted to TMLR](#)

📅 2024

- Used Conformal prediction to interpret Out-of-distribution scores and provide probabilistic guarantees and used OOD scores to construct conformal prediction sets.

Robust One-Class Classification with Signed Distance Function using 1-Lipschitz Neural Networks

[Conference paper - ICML 2023](#)

📅 2023

- Used 1-Lipschitz networks to learn signed distance function from a data point cloud. The algorithm allows One-class classification with robustness guarantees and Implicit Neural Representation.

Making Sense of Dependence: Efficient Black-box Explanations Using Dependence Measure

[Conference paper - NeurIPS 2022](#)

📅 2022

- Attribution method (XAI) by measuring the dependence between regions of an input image and the output. Application in Image Classification and Object Detection.

Accelerating Hypersonic Reentry Simulations using Deep Learning-based Hybridization (with guarantees)

[Journal paper - Journal of Computational Physics](#)

📅 2022

- Replacing costly part of strongly coupled simulation code by neural networks.
- Guarantees for the accuracy of the obtained hybrid code.

SIDE EXPERIENCE

Tutorial teacher (khôlles) in Mathematics

[Lycée Chaptal - CPGE](#)

📅 2015 - 2016

📍 Paris, France

Associative Experience

Project Manager

[Télécom Etude](#)

📅 2015 - 2016

📍 Paris, France

- Telecom's Junior Entreprise: managed projects conducted by Telecom ParisTech students for start-ups.

President

[La Scène](#)

📅 2015 - 2016

📍 Paris, France

- Regular organization of averagely 250 people concerts.

Internships

Data Scientist

[Intento>](#)

📅 2017

📍 Bordeaux, France

Business Developer

[nFinite](#)

📅 2017

📍 Paris, France

Consultant

[Tera Consultants](#)

📅 2016

📍 Paris, France

Goal-oriented Sensitivity Analysis of Hyperparameters in Deep Learning

Journal paper - Journal of Scientific Computing

📅 2021

- Adapting Hilbert-Schmidt Independence Criterion (HSIC) to measure the relative importance of hyperparameters and improve hyperparameter optimization.
- Intensive hyperparameter optimization of MLPs and Convolutional Networks on 400 GPUs using Slurm.

Leveraging Local Variation in Data: Sampling and Weighting Schemes for Supervised Deep Learning

Journal paper - Journal of Machine Learning for Modeling and Computing

📅 2020

- Improving the accuracy of Deep Neural Networks by weighting the samples of the training data set.

Taylor Based Sampling Scheme for Machine Learning

Workshop paper - NeurIPS 2019

📅 2019

- Methodology for sampling efficient designs of experiments for Machine Learning in Computational Physics.
- Accepted in Second Workshop on Machine Learning and the Physical Sciences (NeurIPS 2019).

Deep Unsupervised Representation Learning

MSc Thesis

📅 2018

- Unsupervised representations of the visual concept of counting from images using Autoencoders and Generative Adversarial Networks (GAN)s.
- Evaluation of randomness in artificial dynamics such as cellular automata using Recurrent Neural Networks (RNN)s.