

Avrile Floro

Bures-sur-Yvette, France • floro.avrile@gmail.com
Homepage: avrilemay.github.io • GitHub: github.com/avrilemay

EDUCATION

Institut Polytechnique de Paris, France Sep. 2025 – Jul. 2027 (expected)
PhD Track Master in Data & Artificial Intelligence (jointly offered by École Polytechnique and Télécom Paris)
A Master's program with a reinforced research component: individualized supervision by a faculty advisor and integration into a research lab from the first year. Advisor: Prof. Nils Holzenberger.
Ranked 1st out of 30 in 2025–2026 (overall average: 19.32/20).

Université Paris 8 Vincennes – Saint-Denis, France Sep. 2022 – Jul. 2025
Bachelor's Degree in Computer Science, summa cum laude (highest honors)

Université d'Artois, France Sep. 2022 – Jul. 2023
Master's Degree in Teaching French as a Foreign, Second, or Specific-Purpose Language, cum laude

Université Paris-Dauphine, France Sep. 2016 – Jul. 2017
Master's Degree in Financial Law, cum laude
Master's thesis: *Whistleblower Status and the Scope of Its Protective Provisions*.

Duke University, USA Aug. 2015 – May 2016
Master of Laws (LL.M.) with Certificate in Business Law, GPA: 3.226/4

Université Paris-Sud (now Université Paris-Saclay), France Sep. 2012 – Jul. 2015
Bachelor's Degree in Law; University Diploma in Comparative and International Legal Studies, cum laude

PUBLICATIONS

- [1] *Where Experts Disagree, Models Fail: Detecting Implicit Legal Citations in French Court Decisions*
A. Floro, T. Dhorasoo, S. Pellez, & N. Holzenberger. [arXiv:2603.22973](https://arxiv.org/abs/2603.22973) [Under review]
- [2] *Cultural Binding Heads in Language Models* [BlackboxNLP @ EMNLP 2026]
A. Floro & L. Benedetto. BlackboxNLP 2026, archival track (EMNLP workshop). [arXiv:2605.28543](https://arxiv.org/abs/2605.28543)
- [3] *A Grammar of French Civil Cassation: A Ground-Level Dataset and Tool* [Work in progress]
A. Floro, T. Dhorasoo, N. Holzenberger, T. Le Goff, T. Viard, & M. Boritchev.

RESEARCH EXPERIENCE

Student Researcher, Interpretability Auditing for Bias Measurement Jun. 2026 – Present
Collaboration with Prof. Ali Emami, Emory University
Joint research on the reliability of interpretability methods for social-bias auditing: designing evaluation protocols to establish when readouts of a model's internal representations can be trusted for bias measurement, across multiple open-weight LLMs.

Research Intern, Argument Mining on French Supreme Court Rulings May 2026 – Aug. 2026
i3 – Interdisciplinary Institute for Innovation, CNRS UMR 9217, Palaiseau, France | Supervisors:
Prof. Tiphaine Viard, Prof. Maria Boritchev, Prof. Thomas Le Goff
Computational analysis of civil rulings from the French Supreme Court (*Cour de cassation*), using Judilibre open data. Built a large-scale extraction pipeline exploiting the codified rhetorical structure of *cassation* rulings to map each ground of appeal to the court's response, yielding 121K+ argument

pairs across 86K decisions and the five civil chambers. Lead author of the resulting dataset and tool paper [3].

Student Researcher, Mechanistic Interpretability of Cultural Binding in LLMs Feb. 2026 – Jul. 2026

SAMOVAR, Télécom SudParis | Supervisor: Prof. Luca Benedetto

Lead author of *Cultural Binding Heads in Language Models* [2]. Causal analysis (edge knockout, dose-response α -scaling, generation-time steering) of why LLMs default to equal treatment across cultural groups even when context warrants differentiation. Identified 2–3 mid-layer attention heads mediating cultural binding across 8 LLMs, and showed that models know $3\text{--}5\times$ more than they act upon.

Student Researcher, Implicit Legal Citation Detection Apr. 2025 – Jul. 2026

DIG Team, LTCI, Télécom Paris | Supervisor: Prof. Nils Holzenberger

Lead author of *Where Experts Disagree, Models Fail* [1]. Built a benchmark of 1,015 passage-article pairs annotated by three legal experts ($\kappa = 0.33$) and showed that annotator disagreement is the strongest predictor of failure across all evaluated models: 68% of false positives fall on the 33% of disputed cases. Designed and implemented the full pipeline: Judilibre data collection, JuriBERT bi-encoder training, adversarial filtering, and supervision of the expert annotation campaign.

HONORS & AWARDS

IP Paris PhD Track Excellence Scholarship (2025–2026, renewable), competitive merit-based scholarship from Institut Polytechnique de Paris: €10,000/year cost-of-living grant for the first and second years of the Master's, plus subsidized tuition.

1st place / 87 teams, Influencer or Observer: Predicting Social Roles, Kaggle competition for the Deep Learning course at École Polytechnique, Fall 2025 (see Projects).

PROJECTS

Influencer or Observer: Predicting Social Roles on Twitter – winning entry (88.0% accuracy, 1st of 87 teams) in the Kaggle challenge of the Deep Learning course at École Polytechnique, with Falguny Barua Ema and Saurabh Mishra: stacking ensemble combining LightGBM over 1,373 behavioral features, a fine-tuned CamemBERTa-v2, and a HistGradientBoosting meta-learner. [Repository](#).

Collective Nudging that Scales: How Many Dogs Do I Need to Herd Sheep? – NetLogo agent-based study of how herding effort scales with flock size, with Prof. Ada Diaconescu (LTCI, Télécom Paris); proposes a local sheep-density gradient as a tractable proxy for the flock's centre of mass. [Repository](#).

French Electricity Consumption Forecasting – interactive R Shiny app for rolling forecasts of national electricity demand, with built-in model performance testing. [Live demo](#).

Self-Organizing Map Classifier – Kohonen self-organizing map implemented from scratch in C, used as an unsupervised classifier. [Repository](#).

TEACHING EXPERIENCE

Academic Tutor, Computer Science Bachelor's Program Nov. 2024 – Jul. 2026

Université Paris 8 | 218 hours/year | Appointed for 2024/2025, renewed for 2025/2026

Deliver individualized academic and methodological support across the undergraduate Computer Science curriculum; serve as the primary liaison between faculty and students.

French Teacher

Aug. 2020 – Aug. 2026

The French Class, San Francisco, USA

Six years of one-on-one and group instruction across all CEFR levels (A1 to C2); design personalized curricula and weekly teaching materials tailored to each learner; built an automated note-organization tool to streamline lesson follow-up.

WORK EXPERIENCE

Associate Founder

Apr. 2019 – May 2020

Bonjour SAS, Bogotá, Colombia

Co-launched a social enterprise selling fair-trade cocoa products sourced from Colombian coca-to-cacao crop substitution programs; handled business planning, digital strategy, and retail partnerships.

Divemaster (PADI)

Mar. 2018 – Jul. 2018

Next Level Scuba, Indonesia

Banking Risk & Compliance Consultant (Apprenticeship)

Sep. 2016 – Jul. 2017

PricewaterhouseCoopers (PwC), Neuilly-sur-Seine, France

Audited the compliance procedures of global credit institutions; drafted and revised procedures in anticipation of legislative changes; implemented recommendations of the ACPR (French banking and insurance supervisor), governance, and industry best practices.

TECHNICAL SKILLS & LANGUAGES

Programming: Python (proficient), R, C, Java, Lisp

ML / NLP stack: PyTorch, Hugging Face Transformers, scikit-learn, pandas, NumPy

Other tools: R Shiny, NetLogo, Git, $\LaTeX$

Languages: French (native), English (C1–C2), Spanish (B2)