

# Davide Baldelli

✉ [davide.baldelli@mila.quebec](mailto:davide.baldelli@mila.quebec) | 🏠 [dundalia.github.io](https://dundalia.github.io) | 🐙 [github](https://github.com) | 🔗 [linkedin](https://www.linkedin.com/in/davidebaldelli)

## EDUCATION

### MILA and Polytechnique Montréal

PhD in Computer Engineering

Montréal, Canada

Jan 2025 - Current

- GPA: 4.0
- **Supervisors:** Sarath Chandar, Amal Zouaq.
- **Courses:** Representation Learning, Reinforcement Learning.

### University of Bologna

MSc in Artificial Intelligence

Bologna, Italy

Sept 2021 - Oct 2023

- Grade: 110/110 cum laude.
- **Courses:** Programming for Data Science, Knowledge Representation and Reasoning, Statistical Theory and Methods, Machine Learning, Deep Learning, Computer Vision, Natural Language Processing.

### University of Florence

BSc in Mathematics

Florence, Italy

Mar 2017 - Mar 2021

- Grade: 110/110.
- **Courses:** Mathematical Analysis, Linear Algebra, Group Theory, Geometry, Probability and Statistics, Logic and Computability.

## RESEARCH EXPERIENCE

### National Institute of Informatics

MSc Thesis - International Internship Program

Tokyo, Japan

Mar 2023 - Aug 2023

- **Supervisors:** Akiko Aizawa (National Institute of Informatics), Paolo Torrioni (University of Bologna)
- **Title:** A TWO-step LLM Augmented distillation method for passage Reranking

### University of Florence

BSc Thesis

Florence, Italy

Sept 2020 - Mar 2021

- **Supervisor:** Francesca Romana Nardi (University of Florence)
- **Title** (translated): Big Deviations for Markov Chains

## PUBLICATIONS

### PEER-REVIEWED CONFERENCE PROCEEDINGS

**Baldelli, Davide**, Junfeng Jiang, Akiko Aizawa, and Paolo Torrioni (2024). "TWOLAR: A TWO-Step LLM-Augmented Distillation Method for Passage Reranking". *Advances in Information Retrieval*. Vol. 14608. Lecture Notes in Computer Science. Springer, pp. 465–481. DOI: 10.1007/978-3-031-56027-9\_29. URL: [https://doi.org/10.1007/978-3-031-56027-9\\_29](https://doi.org/10.1007/978-3-031-56027-9_29).

### PRE-PRINTS AND WORKING PAPERS

Govindarajan, Prashant, **Baldelli, Davide**, Jay Pathak, Quentin Fournier, and Sarath Chandar (2025). *CADmium: Fine-Tuning Code Language Models for Text-Driven Sequential CAD Design*. arXiv: 2507.09792. URL: <https://arxiv.org/abs/2507.09792>.

## PROFESSIONAL EXPERIENCE

### MILA and Polytechnique Montréal

Teaching Assistant - Summer High School Internships

Montréal, Canada

Summer 2025

- Designed and delivered educational workshops on Machine Learning fundamentals and Probability & Statistics for high school students.
- Mentored students in hands-on learning activities, fostering interest in AI research.

### Independent Consultant (Upwork Platform)

Freelance Machine Learning Engineer

Remote

June 2024 - Jan 2025

- Delivered a variety of machine learning projects for international clients, consistently earning 5-star ratings for quality and speed.

### Loop AI Group

Machine Learning Engineer & Master's Thesis Co-Supervisor

Full Time - Fully Remote

Sept 2023 - Dec 2024

- Co-supervised a master's thesis titled "Self Assessment Tool for Medical Exam Candidates" in collaboration with the University of Padova. Provided mentorship throughout the entire research process, from problem formulation to final defense. Guided the candidate in literature review, experimental design, machine learning model selection, and evaluation methodology.
- Conducted deep learning project for musical Deepfake detection achieving 94% of F1 score on a proprietary deepfake music dataset.
- Implemented BERT-based email classifier tool for an Italian University's administrative office, achieving 98% accuracy.

## Loop AI Group

Data Scientist Consultant

*Part Time - Fully Remote*

*Sept 2022 - Sept 2023*

- Implemented closed-domain generative question-answering systems on technical topics, exploiting SoTA Pretrained Seq2Seq Language Models.
- Implemented rule-based and transformers-based text retrievers.

## LEADERSHIP & COMMUNITY INVOLVEMENT

---

### Belle Parole Arci APS

Leadership in Non-Profit Organization

*Florence, Italy*

*Nov 2023 - Present*

- Involved in founding and leading Belle Parole Arci APS, a non-profit organization.
- Responsible for organizing community events and managing bar operations as part of the organization's activities.
- Built the website of the organization: **belleparoleaps.it**.

## OTHER ACHIEVEMENTS

---

### AWARDS & RECOGNITIONS

Jun 2025 **Selected Participant**, Y Combinator AI Startup School

*San Francisco, CA*

Jun 2022 **4th place**, Loop Q Prize Competition

*Italy*

### CERTIFICATIONS

Apr 2025 **Agents Course**, Hugging Face

*Remote*

## SKILLS

---

**Programming** Python (fluent); HTML, CSS, R, MatLab, Minizinc, NetLogo (familiar).

**Data Science** Numpy, Pandas, Scikit-learn, Matplotlib, Seaborn, Plotly

**Deep Learning** Pytorch, Pytorch-lightning, Tensorflow.

**NLP** HuggingFace, Transformers, Prompt Engineer, OpenAI GPT.

**Miscellaneous** Linux, Shell,  $\LaTeX$ (Overleaf/R Markdown), Git, Docker, Figma, Blender.

## LANGUAGES

---

**English** Professional proficiency

**Italian** Native proficiency

**References available upon request.**